

特色专题

“物-人-场”交互视域下的数字孪生技术研究进展

游佳莉¹, 邓佳文^{1*}, 焦子韵¹, 罗阿理², 宋轶晗², 邱波³, 任福继^{1*}

摘要 随着新一代信息技术的融合创新,数字孪生(digital twin, DT)技术作为推动数字化转型与智能系统演进的关键支撑,已广泛应用于工业制造、智慧城市、智能交通等领域。然而,现有研究主要聚焦于“物”的建模和分析,较少系统整合“人”和“环境”因素,难以满足复杂智能系统对多层次、全方位交互的发展需求。鉴于此,引入“物-人-场”交互视域,从智能物理实体(物)、智能个体(人)以及虚实融合环境(场)3大核心维度,全面且系统地剖析数字孪生技术的研究前沿与进展。首先,分析以“物”为核心的传统数字孪生技术体系,重点介绍其理论起源、框架及应用等。其次,探讨 AI(人工智能)驱动下的数字人定义、发展脉络、国家政策和核心技术等。最后,将视野拓展至“场”的维度,探究元宇宙多元场景中“场”的应用实践,深入研讨“物”“人”和“场”3个元素的深度融合交互机制,揭示三者如何相互作用、相互促进,为元宇宙的构建提供支撑。进一步,围绕“物-人-场”面临的关键挑战,提出3方面研究建议:(1)构建智能化且具备多层架构的数据融合框架;(2)研究 AIGC(生成式人工智能)赋能的智能虚实映射与虚拟原生进化;(3)构建新的虚拟经济与智能化治理体系。研究成果为构建具备多主体协同感知、多模态智能交互与虚实融合闭环反馈的新型数字孪生系统提供了理论基础与实践启示。

关键词 数字孪生;物-人-场交互;数字人;虚实融合;元宇宙

新一代信息技术的集群式突破与融合创新,促使数字孪生(digital twin, DT)技术迈入变革的新阶段。该技术作为数字化转型进程中的关键驱动力,其发展历程呈现出显著的阶段性演化的特点。它的发展历程可追溯到美国国家航空航天局(National Aeronautics and Space Administration, NASA)航天器维护的初步实践,随后在工业 4.0 时代实现大规模应用,再到当前智慧城市、智能制造等领域的深度融合,这一

技术发展历程展现了持续演进的轨迹。特别值得关注的是,人工智能、物联网、5G(第 5 代移动通信技术)、云计算与边缘计算的协同发展,推动数字孪生技术实现了从“静态建模”到“动态交互”、从“单向映射”到“自主进化”的实质性突破,这些变化正在重塑数字经济时代的技术生态体系。

然而,传统以“物”为核心的建模范式已难以契合复杂系统智能化演进的需求。这一局限主要体现在以下 3 个方面:一是传统方法侧重于物理实体的数字化转型,缺乏对“人”这个关键要素的整合,导致场景和物类型单一,从而限制了交互维度;二是物理实体的数字化形式难以应对动态变化的复杂场景,导致对

1. 电子科技大学计算机科学与技术学院,成都 611731

2. 中国科学院国家天文台,北京 100101

3. 北京科技大学智能科学与技术学院,北京 100083

收稿日期:2025-04-02;修回日期:2025-08-01

作者简介:游佳莉,博士研究生,研究方向为情感分析、大模型技术等,电子信箱:YJL1015004@163.com;邓佳文(通信作者),助理研究员,研究方向为人工智能、大模型技术、情感计算等,电子信箱:dengjiw@uestc.edu.cn;任福继(共同通信作者),教授,日本工程院院士、欧盟科学院院士、俄罗斯工程院外籍院士,研究方向为先进智能、情感计算、智能机器人等,电子信箱:renfuji@uestc.edu.cn

引用格式:游佳莉,邓佳文,焦子韵,等.“物-人-场”交互视域下的数字孪生技术研究进展[J].科技导报,2025,43(20):62-84;doi:10.3981/j.issn.1000-7857.2025.04.00057

环境适应性不足;三是智能协同水平有限,制约了智能系统的响应速度和效率。针对这些挑战,研究人员正逐步将研究视角拓展到“物-人-场”的多层次交互框架。其中,“物”指的是智能物理实体,如工业设备、智能终端和基础设施;“人”涵盖数字人、AI(人工智能)代理及人机协作系统,旨在增强数字孪生的认知与决策能力;“场”代表虚实融合的智能环境,如工业元宇宙、智慧城市及沉浸式数字空间。为了进一步介绍“物”“人”和“场”的差异和共性,表 1 在定义、核心技术等 7 个维度上进行了详细对比。

表 1 “物-人-场”交互视域下数字孪生技术的异同分析

对比维度	“物”——传统数字孪生	“人”——AI驱动的数字人	“场”——物-人融合到元宇宙生态
定义	物理实体的数字化映射	具备智能交互能力的虚拟人	由物理实体的数字孪生和数字人构建的沉浸式虚拟环境
核心技术	建模、渲染、仿真、物联网	形象建模、语言和多维度表达	区块链技术、交互技术、人工智能技术、物联网技术等
数据来源	传感器数据、仿真数据	语音、图像、文本数据	物理世界数据、用户交互数据
交互方式	远程监控、自动优化	语音、表情、肢体动作	多人协作、沉浸式体验
应用场景	智能制造、智慧城市	虚拟主播、智能客服	元宇宙社交、虚拟经济
挑战	数据同步、计算资源	AI情感理解、实时渲染	计算压力、经济治理
发展趋势	高精度实时孪生	更拟人化、智能化	跨平台融合、增强现实

以“数字孪生”技术为核心主线,梳理其以“物”为建模起点,逐步融合“人”的智能交互能力,最终演进至“场”的虚实融合环境的技术发展路径,提出了“物-人-场”交互视域下的数字孪生技术研究进展。相较于传统以物理实体为中心的单一建模范式,本文强调在新一代信息技术的驱动下,数字孪生正逐步演变为支撑多主体感知、多模态交互与虚实融合协同的新型系统的关键技术。在此基础上,系统地分析了“物”“人”和“场”3 类要素的概念、研究现状与典型应用,并进一步阐述了三者之间的关联和数据流结构。

基于“物-人-场”交互视域,数字孪生技术正演化为一个动态开放、智能协同的虚实融合体系,能够支持多主体交互、环境自适应优化及智能服务升级。这

一框架的提出,不仅有助于突破传统数字孪生在复杂场景中的局限性,还能推动数字人、AI 与数字孪生的深度协同,进一步赋能工业智能化、智慧城市及元宇宙等新兴领域,为未来智能社会的构建提供技术支撑。

1 “物”:传统数字孪生的起点

1.1 数字孪生的理论起源和显著特征

数字孪生的概念可追溯到航天军工领域,其演化脉络可以划分为“技术探索、概念提出、应用萌芽、行业渗透”4 个主要阶段,如图 1 所示^[1]。

早在 1969 年,NASA 就在阿波罗计划中首次应用了数字孪生技术,通过建立航天飞行器的虚拟孪生

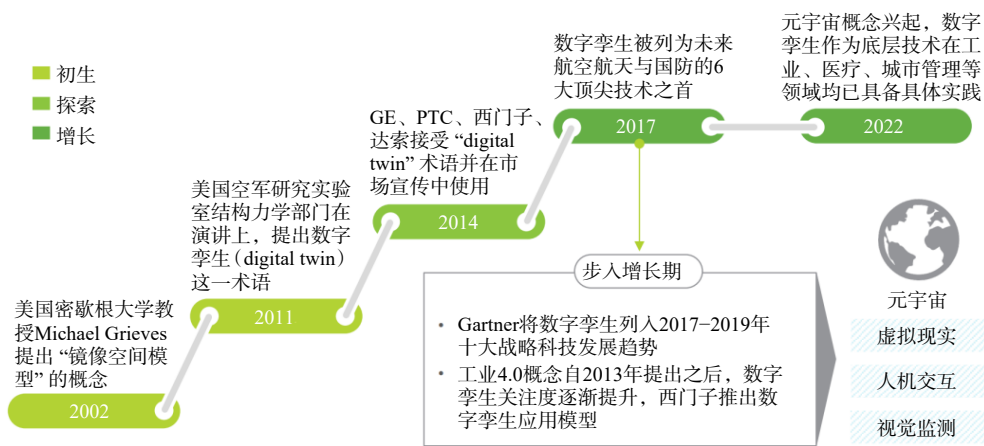


图 1 数字孪生发展历程

体,实现对在轨运行状态的实时监控,并支持紧急情况的应对策略^[2];1991年,Gelernter^[3]提出了“镜像世界”这一类似概念,即软件模型对现实世界进行模拟;2002年,Grieves等^[4]正式提出了“数字孪生”概念,并构建了一个包含实空间、虚拟空间以及两者之间数据/数据流的链接机制的“镜像空间模型”,用以阐述物理实体与虚拟副本之间的交互关系,如图2所示。该模型突出了数字孪生与产品生命周期管理(product lifecycle management, PLM)数据的关联性,为两者的融合奠定了基础。随着物联网、云计算及人工智能技术的快速发展,数字孪生在2010年代中期受到广泛关注,并逐渐从静态仿真模型演进为动态、实时、可自适应的智能系统^[5]。工业4.0的推进使得数字孪生项目,如Verl(volcano engine)得以普及。2017年,数字孪生被美国洛克希德·马丁公司列为未来航天航空与国防的6大顶尖技术之首。2019年,Tao等^[6]对数字孪生进行了系统性定义,强调其不仅包含数字模型本身,还涵盖数据管理、模型更新与服务交付3大核心要素。2021年,ISO 23247标准进一步将数字孪生制造的框架定义为物理制造元素(如人员、产品、资产和流程定义)的虚拟表示。2022年,元宇宙概念兴起,数字孪生作为底层技术在多个领域进行了实践应用,这一演进过程标志着数字孪生从理论探索走向实际应用,并逐步渗透至各行各业^[7]。

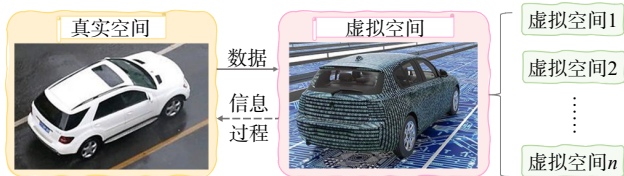


图2 数字孪生示例

数字孪生目前没有统一的理论体系,自其诞生以来,各研究与应用领域提出了很多定义^[8]。为了系统深刻理解数字孪生的概念并深入解析其核心能力^[9],采用文献计量学方法,对近5年Web of Science核心集中收录相关文献进行了系统性分析。基于文献分析结果,研究发现数字孪生技术相较于传统数字化技术,展现出以下4个显著特征。

1) 双向动态映射:数字孪生技术强调虚拟空间中构建物理实体的数字化表征,实现物理实体与其数字模型之间的双向映射、数据互联与状态交互。换言之,

物理实体的各种指标能够真实地呈现在数字孪生体中,而数字孪生体的变化也能够真实反映物理实体的变化。

2) 实时同步检测:通过多源传感数据的实时采集,数字模型能够精确、动态地反映物理实体的状态变化,包括外观形态、性能参数、空间位置及异常情况等等。

3) 共生共智:在“共生”层面,这一特性表现为数字孪生与物理实体在系统整个生命周期内的同步创建和协同发展,二者之间建立了相互依赖、相互推动的共生联系。而“智能”方面则涉及2个层面:在微观尺度上,指的是数字孪生系统内部各模块间的智能资源共享,涵盖数据交换与算法协同;在宏观尺度上,则是在多个数字孪生系统组合形成的复杂系统中,各子系统间智慧因素的深度整合与协同优化,进而提升了系统的整体智能决策水平。

4) 闭环优化控制策略:数字孪生技术的关键效用体现在对物理实体内部机制的揭示、运行规律的分析以及未来趋势的预测,进而通过仿真手段制定优化策略,实现对物理实体的闭环式控制与持续优化。

1.2 国家重大需求

中国对数字孪生技术的战略意义予以高度重视,自“十三五”规划便开始部署相关政策,并在智能制造、智慧城市建设、数字化政府等领域推动了该技术的试点与示范应用。步入“十四五”规划阶段,伴随《“十四五”数字经济发展规划》的发布,数字孪生技术被明确定位为支撑数字经济高质量发展的关键技术之一,标志着技术发展进入了加速阶段。如表2所示,政策支持力度不断强化,产业生态系统逐渐成熟,应用领域持续拓展,促进了数字孪生在多行业的融合。

自2023年起,中央和地方政府相继出台了一系列政策文件,旨在支持数字孪生技术的研发、标准化建设及产业化应用。中央政府层面,国家发展改革委、工业和信息化部等部门发布多项指导性政策,明确数字孪生在智能制造、城市治理、交通管理、能源调度等方面的核心作用;地方政府层面,各省市积极响应,结合区域特色出台专项政策,加快数字孪生技术的试点推广,打造地方示范工程。例如,北京、上海、深圳等地在智慧城市和智能交通领域率先布局,赋能城市治理现代化;长三角、粤港澳大湾区等区域

表 2 2020—2025 年部分数字孪生相关政策概览

年份	发文单位及政策名称	主要内容
2020-04	国家发展改革委,《关于推进“上云用数赋智”行动培育新经济发展实施方案》的通知	开展数字孪生创新计划。鼓励研究机构、产业联盟举办形式多样的创新活动,围绕解决企业数字化转型所面临数字基础设施、通用软件和应用场景等难题,聚焦数字孪生体专业化分工中的难点和痛点,引导各方参与提出数字孪生的解决方案
2020-12	工业和信息化部,《工业互联网创新发展行动计划(2021-2023年)》	支持建设云仿真、数字孪生等技术专业型平台,加快信息技术创新应用,研发构建数字孪生创新工具,打造创新解决方案,发展基于数字孪生技术的工业智能解决方案
2021-06	国务院,《“十四五”数字经济发展规划》	提升城市综合管理服务能力,完善城市信息模型平台和运行管理服务平台,因地制宜构建数字孪生城市
2021-11	国务院,《“十四五”大数据产业发展规划》	融合应用市场加速繁荣,场景挖掘将从边缘浅层向核心深层拓展。基于数字孪生体的制造执行、跨系统跨产业链的综合性分析等深层次应用将加速涌现
2022-01	国家发展改革委、水利部,《“十四五”水安全保障规划》	要推进数字流域、数字孪生流域建设,实现防洪调度、水资源管理与调配、水生态过程调节等功能,推动构建水安全模拟分析模型,要在重点防洪区域开展数字孪生流域试点建设
2022-10	国家能源局,《能源碳达峰碳中和标准化提升行动计划》	加快完善能源产业链数字化相关技术标准体系,推进能源各领域数字孪生、能源大数据、智能化等技术标准制修订
2023-01	中共中央、国务院,《数字中国建设整体布局规划》	推动生态环境智慧治理,加快构建智慧高效的生态环境信息化体系,运用数字技术推动山水林田湖草沙一体化保护和系统治理,完善自然资源三维立体“一张图”和国土空间基础信息平台,构建以数字孪生流域为核心的智慧水利体系
2024-01	中共中央办公厅、国务院办公厅,《浦东新区综合改革试点实施方案(2023-2027年)》	创新高效能城市治理体系。建立城市治理智能化统筹提升机制,健全城市信息模型(CIM)平台,推动智慧城市时空大数据平台应用,打造数字孪生城市
2024-11	国务院,《水利部关于推进水库、水闸、蓄滞洪区运行管理数字孪生的指导意见》	到2027年,推进具有防洪任务的已建大型及防洪重点中型水库、大型水闸、国家蓄滞洪区数字孪生建设,迭代优化数字孪生水利工程先行先试建设成果,促进数字孪生工程与数字孪生流域、数字孪生水网互联互通与信息共享,初步建成运行管理智能应用体系

则依托先进制造业基础,推动数字孪生赋能工业互联网和智能制造,促进数字经济发展。

政策的持续推进不仅加快了数字孪生技术的应用实践,还显著提升了行业规范化水平。国家层面逐步建立健全技术标准、数据管理规范和安全管理体系,形成了较为完备的政策支持体系,为企业提供了明确的发展方向和技术导向,有效推动了产业链的协同创新。同时,产业资本加速投入,企业纷纷增加在数字孪生领域的研发投入,进一步推动了技术升级和商业模式的创新。综上,在国家政策的强力推动下,中国数字孪生技术正处于前所未有的发展高峰,产业生态系统日益完善,未来将在智能制造、智慧城市、虚拟仿真、数字经济等多个领域发挥更加关键的作用,为中国数字化转型和高质量发展提供坚实支撑。

1.3 数字孪生系统架构

数字孪生核心在于将物理世界的数据引入数据层,通过构建精准的数字孪生体,模拟、分析和优化现实世界实体的全生命周期^[10]。图 3 展示一个完整的数

字孪生系统框架,从底层的物理实体到顶层的应用场景,形成了一个从数字采集到智能应用的完整闭环。当前研究普遍遵循一个 4 层架构:数据层、功能层、建模层、应用层以及贯穿全流程的数据保护与伦理保障的隐私安全层形成了一个从数据采集到智能应用的完整闭环。

1) 数据层:可分为数据采集和数据融合 2 个子层,是数字孪生的基础。其中,数据采集层汇聚多源异构数据,包括可穿戴设备(心率、皮肤电反应等生理信号)、视觉传感(摄像头、动作捕捉)、语音与文本输入,以及环境上下文信息。典型平台如 MetaHuman、OpenPose 等已支持高精度实时动作与面部表情捕捉;数据融合层通过多模态深度学习技术实现异构数据的同步处理与特征融合。主流方法包括跨模态注意力机制与图神经网络可同时捕捉时序关联与空间语义信息。

2) 功能层:为数字孪生提供通用技术支撑,通常包含 4 大模块,包含仿真引擎、可视化渲染、实时控

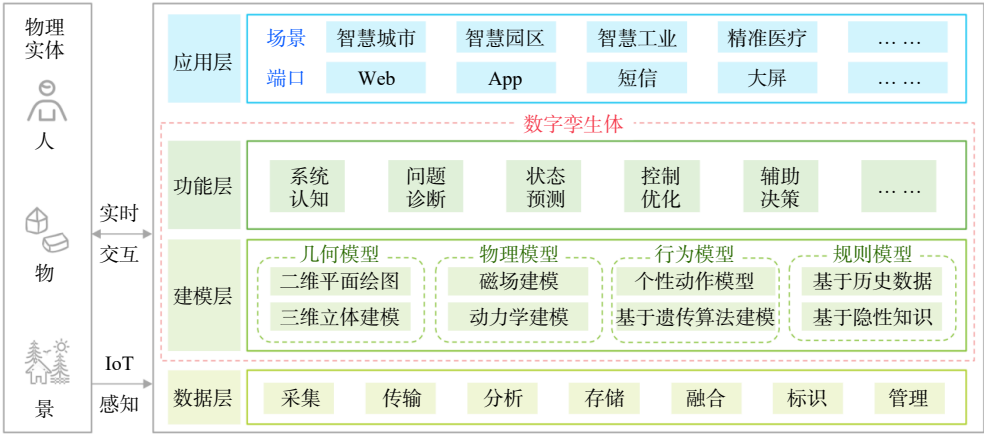


图3 数字孪生系统框架

制和分析优化。其中,仿真引擎通常是基于物理模型或数据驱动模型;可视化渲染通过多种方法实现数字模型的逼真视觉效果,主要包含基于物理的渲染、实时渲染和云渲染;实时控制通过数字线程反馈指令到物理实体,如自动驾驶孪生中的路径规划;分析优化集成 AI 算法(强化学习、因果推理)进行预测性维护或资源调度。GE Predix 平台通过功能层实现航空发动机寿命预测是一个典型应用。

3) 建模层:是区分数字孪生与简单仿真的关键,需解决多尺度与多物理场耦合问题。换言之,建模层是数字孪生的核心,通常采用多种建模技术创建物理实体的数字模型,主要包括 3D 扫描、参数化建模和逆向工程等。

4) 应用层:是数字孪生技术具体面向行业需求的关键环节。通过数字孪生的核心技术和行业实际场景深度融合,形成定制化解决方案。

5) 隐私与伦理层:是贯穿全流程的防护网。数字孪生面临数据泄露(如人脸孪生被恶意复制)、模型攻击(对抗样本误导决策)等风险,需分层防护。针对数据安全,通常采用差分隐私处理敏感数据、联邦学习实现“数据不动模型动”;针对模型安全,通常采用对抗训练提升鲁棒性或区块链存证确保模型不可篡改;针对伦理治理,欧盟 AI 法案要求数字孪生符合透明

性准则,同时可以通过虚拟人伦理检测隐私泄露。

1.4 数字孪生关键技术

在数字孪生系统的整体架构中,建模、渲染、仿真和物联网(internet of things, IoT)4 大核心技术相互协同,共同构建一个完整的数字孪生生态系统^[11],如图 4 所示。这 4 个核心技术各司其职又紧密配合,实现了对物理实体的全方位数字化映射、动态可视化呈现以及智能化决策支持。进一步,表 3 介绍了这 4 大技术的主要手段、发展趋势以及主要局限性。

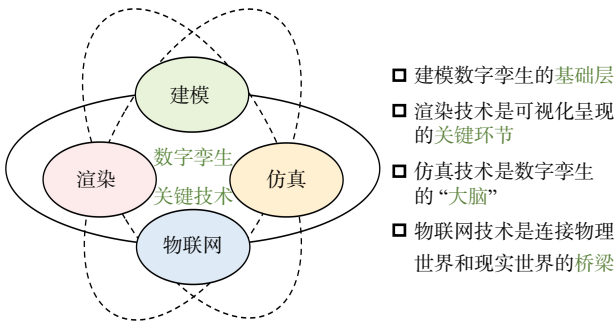


图4 数字孪生关键技术

1) 建模技术作为数字孪生的基础层,主要负责物理对象的数字化表征^[12]。通过几何建模、物理建模和数据驱动建模等多种方法,构建与真实世界高度一致的虚拟模型^[13-14]。这一过程不仅需要精确捕捉物理对象的空间特征,还要充分考虑其内在的物理特性和行

表3 数字孪生4大核心技术解析

技术	关键手段	发展趋势	主要局限性
建模	3D扫描、物理建模、逆向工程、AI建模	参数化、混合建模	复杂系统建模难、计算成本高
渲染	实时渲染、光线追踪、云端渲染	AI优化、5G低延迟	实时性挑战、硬件成本高
仿真	有限元分析、多物理场耦合、实时仿真	闭环仿真、云端协同	计算资源需求大、模型简化误差
IoT	传感器、边缘计算、5G	AIoT、DTaaS	数据安全、异构数据整合难

为规律。

在几何建模中,主要用于构建高精度的3D形态,包括计算机辅助设计(computer-aided design, CAD)、建筑信息建模(building information modeling, BIM)、点云重建等技术。Tao等^[15]系统阐述了CAD和BIM在制造领域数字孪生建模中的应用;Meng等^[16]提出了一种基于点云数据的几何建模方法对无人机进行了数字孪生,可以实现实时系统监控和云处理、VR连接、通过VR技术实现人机交互等;Pan等^[17]提出一种3D深度学习增强空隙增长方法来创建建筑物结合数字孪生。

在物理建模中,要用于模拟真实物理世界的力学、热学、电磁场等属性,典型方法包括有限元分析(FEA)和多体动力学(MBD)。Grieves^[18]详细讨论了多体动力学建模在数字孪生中的应用;Kapteyn等^[19]提出了一个概率图形模型作为数字孪生及其相关物理资产的正式数学表示;《The finite element method: its basis and fundamentals》^[20]中利用有限元分析来模拟结构受力情况。

AI技术加速了数据驱动建模的发展,核心方向包括:其一,深度学习与图神经网络自动化建模,例如,3D点云到CAD模型转换^[21];其二,强化学习+物理引擎,结合游戏引擎(Unity、Unreal Engine)进行仿真建模,提高交互性^[22];其三,生成式AI(GAN、Diffusion Models),用于自动化生成逼真的数字孪生模型,例如,通过扩散模型生成高质量工业组件^[23]。

2) 在可视化过程中,渲染技术扮演着重要的角色^[24],它能够数字模型精确转换为易于理解的可视形式。通过采用实时渲染、光线追踪等前沿技术,得以实现对数字孪生模型的高保真再现,进而为用户带来深入的交互体验。随着图形处理技术的不断提升,当前渲染技术的发展趋势主要体现在混合渲染策略、神经渲染方法的轻量化,多模态数据的实时融合及真实感渲染^[25]。混合渲染架构融合了光栅化与光线追踪技术,并借助深度学习对全局光照计算予以优化,提升了计算效率与渲染精度^[26];轻量化神经渲染(如Instant-NGP)凭借减少网络参数量,达成了低成本、高质量的实时渲染^[27];同时,多模态数据融合技术的应用增强了渲染系统对真实环境的适应性^[28],为数字孪生在智慧城市建设、工业生产模拟以及虚拟现实仿真

等领域的应用提供了坚实的技术支持;在这些研究基础上,真实感渲染技术凭借其基于物理光照模型的理念,模拟光与材质的相互作用,进一步提升了视觉效果的真实度。研究主要集中在以下3个方面:将物理光照(physically based rendering, PBR)理论融入神经辐射场(neural radiance fields, NeRF)框架,联合估计场景几何、材料和照明参数,从而实现真实感渲染与重光照^[29];采用神经网络直接从单张输入图像生成能够适应新视角和光照条件的材质视觉效果,突破了传统渲染路径,支持实时材质变换^[30];结合PBR材质与神经光场架构,通过光采样网络和双向反射分布函数(bidirectional reflectance distribution function, BRDF)表示,创建了一种更为精确的真实感神经渲染器,能够处理复杂材料如皮肤和玉石等^[31]。

3) 仿真技术则是数字孪生的“大脑”,通过物理仿真、离散事件仿真等方法^[32],对系统行为进行预测和优化。它能够模拟各种工况条件下的系统响应,为决策提供科学依据。特别是在结合人工智能技术后,仿真精度和效率都得到了显著提升^[33]。仿真技术在数字孪生中的应用正朝着多尺度融合、高效实时计算和智能增强模拟方向发展。多尺度仿真结合微观物理建模和宏观系统仿真,实现跨层次的精确预测^[34];实时仿真加速技术(如基于深度学习的代理建模)通过训练神经网络来替代复杂的数值计算,大幅降低计算成本^[35];此外,智能增强仿真正在利用强化学习和生成对抗网络优化仿真策略,使数字孪生系统能够在复杂环境中进行自适应优化^[36]。这些技术的进步推动了数字孪生在制造、智慧城市和航空航天等领域的广泛应用,为更精准的预测分析和智能决策提供了强有力的支撑。

4) 物联网技术作为连接物理世界和数字世界的纽带,依托传感器网络、边缘计算等关键技术,实现了数据的实时采集与传递^[37]。这些数据不仅用于模型的初始构建,更是维持数字孪生与物理实体同步更新的关键。随着5G/6G通信技术的发展,物联网在数字孪生系统中的作用愈发重要^[38]。

4大技术相互关联、互为助力,构筑了数字孪生技术的根本支撑结构,助力数字孪生在各个行业的广泛应用发展。

1.5 数字孪生技术的典型应用案例

数字孪生技术作为数字化转型的核心引擎,正从

传统的物理场景建模向更复杂的环境与系统仿真拓展,逐步形成“场景孪生”这一重要应用方向。场景孪生通过高精度建模、实时数据融合和智能分析,实现对物理空间的全息映射和动态演化预测,在智慧城市、智能制造、应急管理等领域展现出广泛的应用价值。

1) 文化保护领域。随着中国传统文化的快速流失,包括艺术品侵蚀和非物质文化遗产传承危机等问题,文化数字化保护已成为中国工程院重点关注的科研方向。在这一背景下,数字孪生技术为文化遗产的数字化保存与活化利用提供了创新解决方案。任福继研究团队以“科技赋能文化传承”为核心理念,系统性地开展了文化遗产数字化保护研究,如图5所示。团队以北魏佛像艺术为研究对象,采用高精度三维扫描技术对佛像进行重建,最终建立数字孪生模型,实现文物的永久性数字化保存;同时运用生成式AI技术对受损佛像进行虚拟修复,重建其历史原貌。在保护实践基础上,团队还探索数字文化的商业模式,例如构建佛像展览元宇宙空间,增加艺术家的商业变现能力,提高文化传承能力,从而形成“保护—修复—传承—创新”的良性循环机制。这一案例的成功实施为后续中国传统文化的保护提供了指导。

2) 天文领域。在人类对宇宙探索的过程中,天文学始终扮演着引领者和开拓者的角色。但是传统的天文学研究主要依赖于高昂的望远镜设备,导致观测资源有限;观测数据规模庞大,导致处理和存储困难;宇宙演化模拟复杂,导致计算成本高。随着数字孪生技术的兴起,为天文学研究提供了全新的工具和方法,如图6所示。例如,对大型天文望远镜进行数字

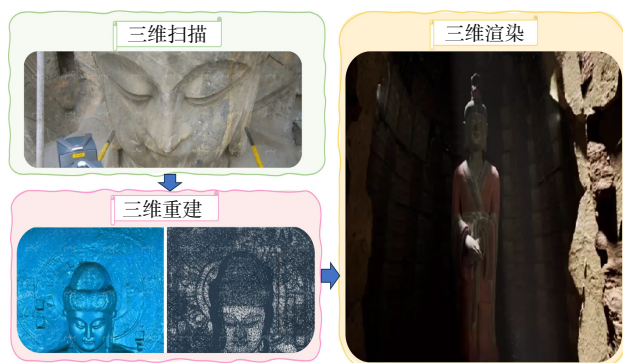


图5 元宇宙-文化保护解决方案框架

孪生管理,利用数字孪生模型对其各个子系统(如镜面、热管理系统等)进行实时监控和性能评估;通过传感器收集数据,结合AI与机器学习算法,实时评估望远镜在轨运行中的状态,预见可能出现的故障并进行预防性维护;其次,数据孪生还通过虚拟与现实的深度融合,为天文学家和普通大众带来了前所未有的体验。例如,创建设备和观测环境的虚拟模型,通过实时数据和物理定律(如牛顿万有引力定律、开普勒定律等)进行轨迹计算和预测;模拟和预测环境因素(如温度、湿度、振动等)对望远镜的影响,从而调整观测策略,确保数据质量。更进一步,可以使用数字孪生模拟宇宙中天体的运动、变化以及宇宙环境的演化,以深入理解宇宙的构成与演化规律。从上述案例不难看出,数字孪生正在为天文学的各个层面提供新的思路,不仅能够帮助天文学家更高效地进行研究,还能将宇宙的奥秘以更加直观和生动的方式呈现给每一个人。

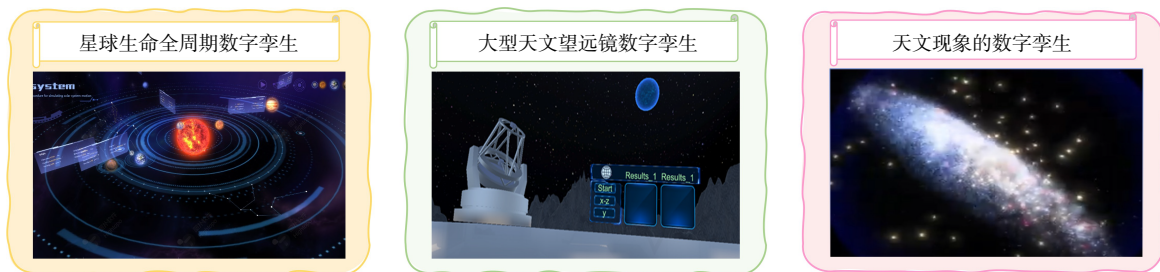


图6 数字孪生和天文领域结合的案例

2 “人”:数字AI驱动的数字人技术发展

2.1 数字人定义与发展脉络

数字人是指通过计算机图形学、人工智能、动作

捕捉等先进技术,实现形象逼真、行为拟人、可实时交互的虚拟人物。数字人既包括高度仿真的虚拟偶像、虚拟主播,也包括能够在特定应用领域实现实时交互的智能数字助手,其中部分数字人可能源自对真

实个体的数字孪生构建。数字人技术的发展主要经历了3个关键阶段。

1) 第1阶段(20世纪80—90年代)是基础计算机图形学(computer graphic, CG)虚拟人物阶段,这一时期的虚拟人技术主要依托于计算机动画和三维图形技术。1982年,日本动画《超时空要塞》首次实现了虚拟偶像的突破,其女主角林明美通过动画技术塑造,成为世界上第1个具有完整人物设定和音乐作品的虚拟歌姬;1984年,英国Channel 4电视台推出的虚拟主持人Max Headroom,虽然主要依靠特效化妆和传统动画技术实现,但开创了虚拟人物在媒体传播领域的先河;进入20世纪90年代后,随着计算机图形技术的进步,1991年电影《Terminator 2》(《终结者2》)中采用CGI技术打造的液态金属机器人T-1000,展现了当时最先进的数字角色特效;2001年上映的《最终幻想:灵魂深处》则首次尝试全CG制作电影人物。这一阶段的虚拟人物虽然在视觉效果上取得突破,但仍存在动作表现机械化、交互功能缺失等技术局限,主要依赖离线渲染实现,单帧渲染时长可达数小时,应用场景主要集中在影视特效领域。

2) 第2阶段(21世纪初—2015年)是动作捕捉与实时渲染阶段,这一阶段以动作捕捉(motion capture, MoCap)技术与实时渲染技术的突破性发展为标志,实现了虚拟角色从静态呈现到动态交互的重大跨越。2001年,英国推出全球首个虚拟新闻主持人阿娜诺娃(Ananova),开创了数字人在媒体领域的应用先河;2004年,中国中央电视台CCTV-6频道推出国内首位虚拟电视节目主持人“小龙”,标志着数字人技术在国内的首次重要实践;2007年,日本Crypton Future Media公司开发的虚拟歌姬(初音未来)问世,其基于VOCALOID语音合成引擎和CG技术的完美结合,成为首个具有全球影响力的虚拟偶像;2009年,詹姆斯·卡梅隆执导的电影《Avatar》(《阿凡达》)将动作捕捉技术推向新高度,其创新的面部表情捕捉系统实现了虚拟角色前所未有的情感表现力。与此同时,Unity、Unreal Engine等实时渲染引擎的日趋成熟,不仅大幅提升了虚拟角色的视觉表现质量,更实现了虚拟角色与用户的实时互动,为数字人技术在游戏中、影视等领域的广泛应用奠定了技术基础。

3) 第3阶段(2016年至今)是AI驱动的智能数

字人阶段。随着深度学习、自然语言处理和计算机视觉技术的突破性发展,数字人实现了从单向展示到智能交互的质的飞跃。2013年,“虚拟邓丽君”技术首次实现与真人歌手周杰伦的隔空对唱,开创了数字人表演的新纪元;2020年是技术爆发的一年,三星NEON在国际消费电子展(consumer electronics show, CES)上展示了自主生成表情和动作的能力;百度推出的虚拟人“小李”实现了跨时空对话;国内相继诞生了虚拟达人Aimee和超写实博主琪拉Qeelia。这一阶段的代表作品还包括微软“小冰”、超写实虚拟人“AYAYI”和国风虚拟偶像“柳夜熙”,它们不仅具备自然交互能力,还能进行情感表达和内容创作,标志着数字人技术正式进入智能化时代。

随着数字人技术的阶段性演进,数字人的应用形态和功能特征呈现出显著的多样性。由于虚拟数字人的核心价值是提供类“人”服务于体验,可以将现有的数字人按照交互能力和视觉风格等维度进行类型划分,如图7所示。

1) 按照是否具备和用户交互的能力,可以将虚拟数字人划分为交互型和非交互型。进一步地,根据驱动方式的不同,交互型数字人可以细分为AI驱动型和真人驱动型。相比于真人驱动型,AI驱动型虚拟数字人依托智能系统,能够自主感知外部信息,进行语义理解和决策,并驱动数字人模型与用户进行自然交互,实现更高层次的自主性和智能化。

2) 根据视觉风格的不同,虚拟数字人可以划分为2D数字人和3D数字人。2D数字人主要通过绘制图像、图像合成以及动画制作等手段,在二维平面内构建图像或动画,因而呈现出一定的平面性,缺少立体感与深度效果。2D数字人适合对交互性和真实性要求不高的场景,如虚拟主播、客户服务、广告推广及教育娱乐等。相较而言,3D数字人技术融合了三维建模、动态捕捉、面部识别、实时渲染以及人工智能算法,创造出基于三维模型的真实3D数字形象,显著提升了其真实感和立体效果。这种数字人模型特别适用于对真实感和交互性要求较高的应用场景,例如虚拟偶像、虚拟客户服务、虚拟导览服务等。

2.2 国家重大需求

随着人工智能和数字孪生技术的快速发展,数字人正逐步渗透至多个领域。作为虚拟与现实深度融



图7 按照交互能力和视觉风格进行数字人划分的框架

合的重要载体,数字人的广泛应用不仅推动了新型数字经济的增长,也对社会治理、文化传播和产业升级产生了深远影响。为了促进数字人技术的健康发展,各国政府和国际组织纷纷出台相关政策支持。

从政策发展历程来看,国内数字人相关政策的出台可追溯至“十三五”时期。国家提出信息化发展战略,并相继发布人工智能技术发展规划、虚拟现实产业发展规划等重要政策,为数字人技术的萌芽奠定了基础。进入“十四五”时期,国家进一步深化“数字经济”发展规划,全面推动社会数字化转型,夯实数字人相关的核心技术体系,为该领域的创新发展提供了良好的政策环境和广阔的未来空间。2021年3月,国家将虚拟数字技术的发展正式纳入“十四五”规划,明确指出虚拟数字技术创新是产业升级和科技强国建设的重要路径之一。在地方层面,2022年8月,北京市经济和信息化局发布了全国首个数字人产业专项支持政策《北京市促进数字人产业创新发展行动计划(2022—2025年)》,提出到2025年北京市数字人产业规模突破500亿元的目标,为数字人产业化发展提供了重要指引。同年10月,科技部与住房和城乡建设部提出研究新型全感知、自然交互、虚拟现实、全息影像等视听技术,进一步推动数字人技术的深化应用。2024年4月,人力资源和社会保障部等9个部门联合发布《加快数字人才培养支撑数字经济发展行动方案(2024—2026年)》,强调数字人才在支撑数字经济发展中的核心作用,旨在培养高水平数字人才,为产业高质量发展赋能蓄力。在国家最高决策层面,2025年3月全国两会期间,全国政协委员刘庆峰提出

应进一步明确机器人分类及应用场景,建议将面向消费端的机器人细分为虚拟数字人与实体机器人,为数字人产业的发展方向提供了更具针对性的政策支持。

总体来看,数字人产业的政策支持体系已从早期的信息化战略逐步发展为覆盖技术研发、产业创新、人才培养和场景应用的全链条支持体系。在国家 and 地方政府的持续推动下,数字人产业正迎来前所未有的发展机遇。

2.3 AI 赋能下数字人的核心技术

数字人的核心能力主要体现在形象建模、语言表达和多维度交互3个关键维度。如图8所示,这3个维度协同作用,使数字人不仅外观逼真,还能与用户建立情感连接,提供个性化服务。在深度学习和大模型的推动下,数字人的视觉逼真度和智能化水平显著提高,目前这3个维度的研究进展如下。

1) 形象建模旨在通过高精度的3D建模、动作捕捉和渲染技术,打造出逼真且富有表现力的虚拟形象,使数字人在外观、表情和动作上与真人高度接近,

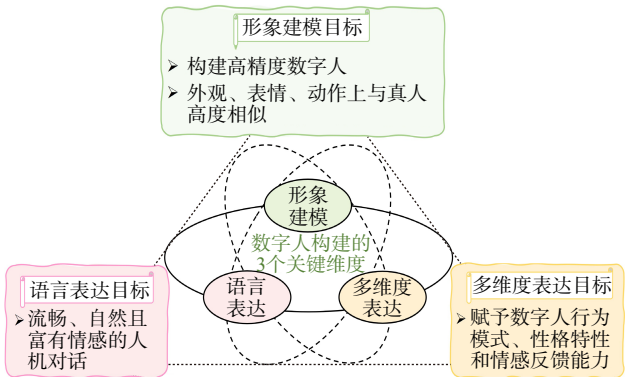


图8 数字人构建的关键要素

甚至突破物理限制,创造出更具艺术性或功能性的视觉体验。同时,在个性化情感交互方面,数字人应综合考虑用户的长期偏好与实时情感状态,通过情感计算和个性化算法,动态调整交互策略,从而提供更加自然、贴心和沉浸式的互动体验。

对于完全虚构的角色,Liu^[39]认为人格是虚拟人格的重要特征。因此,总结了目前计算机人格模型的相关研究,并提出了人格和其他心理变量的形式定义。Cheng 等^[40]基于在现实世界中进行的相关心理学研究,确定了一组可能反映情感数字人个性的服装因素(颜色、设计和类型),并进一步利用“五大”人格模型来评估情感数字人的服装因素对其感知能力的影响。

对于真实人物的数字孪生,传统方法依赖 3D 变形模型 3DMM,此方法通过将 2D 图像重建为 3D 模型以实现虚拟外观建模,但是存在局限性^[41-42]。研究者为解决这个问题,提出了基于部位感知的学习前向蒙皮模块^[43],该模块由模型 SMPL-X 的参数空间驱动,从而实现模型 X-Avatars 的表达动画。随着数字孪生技术的进步,数字人构建已实现从静态建模到动态孪生的跨越。通过融合 3D 扫描、高精度动作捕捉和生物传感器等多源实时数据,可构建与物理实体完全对应的 1:1 高保真数字人模型。同时,借助物联网传感器网络(包括惯性测量单元、光学动作捕捉系统等)的实时数据流,数字孪生系统能够持续更新并动态优化数字人的微表情、肢体动作乃至生理状态参数,实现虚实空间的精准同步与交互。

2) 语言表达旨在通过自然语言处理(natural language processing, NLP)和预训练模型,实现流畅、自然且富有情感的人机对话。通过深度学习和上下文理解,数字人能够准确捕捉语境、表达情感,并在多语言、多场景下提供精准的交互服务。传统对话系统通常依赖情感词典或简单的分类模型(如“积极/消极”标签)生成回应,导致情感表达机械化且缺乏动态适应性,难以满足复杂交互需求。

为了生成逼真的语音,Bark 模型通过端到端的方式生成接近人类水平的语音输出,使数字人具备流畅性、清晰度和情感表现力。语音合成技术(text-to-speech, TTS)^[44]与语音识别系统(automatic speech recognition, ASR)^[45]的结合,使数字人能够准确理解人类语言并做出智能化响应。此外,研究人员开发了多

种算法来提升语音合成的质量,包括基于规则的系统^[46]、统计模型(如隐马尔可夫模型,hidden Markov models, HMMs),以及最新的神经网络方法,如深度神经网络(deep neural networks, DNNs)^[47]和卷积神经网络^[48]。这些神经网络模型通过大规模语音样本的训练,能够提取语音的基本特征,并学习如何将文本转换为连贯的语音波形。在语音合成技术的基础上,为提升数字人声音的个性和情感表达能力,研究者采纳了情绪检测与语音变换技术。该系统通过解析语音的语调、节奏及音量等要素,实现对说话者情绪状态的准确识别,并据此对合成语音进行实时调整,以达到情感匹配的目的。同时,为了适应国际用户群体,当前的语音合成技术实现了对多语言及方言的兼容性。借助跨语言训练与方言转换算法,虚拟数字人得以运用多种语言进行对话,并再现各地口音,以此提供更具亲和力和地域特色的用户交互体验。

传统数字人已实现流畅的语言表达,且大语言模型的突破性发展为数字人语言交互带来了质的飞跃。通过大模型的强大数据处理和学习能力,数字人能够精准捕捉用户偏好,构建高度个性化的交互模型;同时,结合情感计算与心理学理论(如认知重评、共情引导策略)和上下文语义理解,数字人能够实现自然的情感表达与共情交互,显著提升拟人化体验。例如,研究人员开发了新一代共情陪伴大模型 Emoha^[49],模型能够根据用户需求提供针对性的回复方案,善用疑问和引导的方式,和用户共同构建安全温暖的聊天环境,提供深层次的情绪支持;Na^[50]提出了专门为认知行为治疗技术设计的大语言模型 CBT-LLM 模型,在基于认知行为疗法的中文心理健康问答数据集上进行微调,擅长在心理健康支持任务中生成结构化、专业化和高度相关的回复。情感引导多模态对话模型(ELMD)将整体框架分解成情绪检索模块、情绪反应预测模块和情绪增强的反应生成模块^[51],通过两阶段训练策略来理解多模态信息背后的微妙情绪,生成更加细粒度的富有情感的对话。Llanes-Jurado 等^[52]创建了一个基于大语言模型能够进行半引导对话的实时对话数字人,该数字人融合了实时语音合成、实时声唇同步、实时面部表情生成、对话系统,实验证明能够有效地激发人类情绪。

随着数字孪生技术的深度应用,数字人的语言表

达正在经历从机械发声到情感化交互的范式转变。这一演进主要体现在以下 2 个维度: 首先, 在语音生成层面, 数字孪生技术通过整合多维用户数据(包括声纹特征、语调历史记录及情感偏好), 实现了高度个性化的语音合成。以京东 LiveTTS 系统为例, 其基于用户画像的深度学习模型能够动态调整语速、音调和情感倾向, 使数字人的语音输出更具人性化特质。其次, 在多模态协同层面, 数字孪生建模技术构建了语音、唇形与表情数据的动态关联网络。这种跨模态的精准映射确保了数字人在语言表达时, 口型动作、面部表情与语义内容的高度一致性。Synthesia 公司的 AI 视频生成平台正是这一技术的典型代表平台, 其通过实时数据驱动实现了拟真度极高的数字人表达效果。

3) 多维度表达。目标是通过情感计算、个性化算法等技术, 赋予数字人独特的行为模式、性格特质和情感反馈能力, 使其能够在交互中展现出更丰富的层次感和真实感, 实现多模态优化转向跨模态时序与语义一致性建模, 从而为用户提供沉浸式、个性化的体验。在表情表达上, van der Struijk 等^[53]通过面部动作编码系统(facial action coding system, FACS)提取动作单元(action unit, AU)特征, 驱动数字人微表情生成, 并利用轻量化技术(如模型剪枝)适配机器人硬件限制。在肢体表达上, Garg 等^[54]通过时间变换将情感数字人的行走和跑步动作转换为同一时间段, 然后选择合适的融合函数来整合情感数字人的行走和跑步动作, 从而提高了情感数字人在复杂虚拟环境中从起点到事件点的运动保真度。

随着大模型的出现, 大模型推动了多模态技术的深度融合, 使数字人在视觉、听觉、触觉等多维度上实现协同表示, 进一步增强了其在虚拟环境中的真实感和沉浸感。具体而言, 基于大模型融合情感计算、个性化建模与交互规则, 通过注入情感计算理论, 实现细粒度情感语义生成、多维度行为协同决策及安全可控的对话逻辑; 动态适应复杂情感表达、维持人设一致性, 并通过跨模态绑定(语音、表情、动作)驱动拟人化交互, 为虚拟数字人形象构造与行为驱动提供智能化、情感化的决策支持, 形成“感知—理解—表达”闭环系统。

随着数字孪生的演进, 数字人的多维度交互能力

已实现从单向指令到情境感知的质变突破。这一进步主要体现在以下 2 个关键维度: 首先, 在环境感知层面, 数字孪生技术通过构建虚实融合的空间映射系统, 使数字人能够实时感知并理解物理环境状态。以智能制造场景为例, 数字孪生驱动的虚拟质检员可以同步获取生产线传感器数据, 动态调整检测策略和交互方式。其次, 在情感交互维度, 数字孪生系统整合了用户生理数据(如心率、表情变化)与行为特征, 构建了闭环反馈机制。教育领域的应用案例显示, 通过实时分析学习者的微表情和操作数据, 数字教师能够智能调整教学节奏和内容呈现方式, 实现真正的个性化互动体验。

2.4 数字人技术的典型应用案例

在 5G 通信、人工智能、虚拟现实等新一代信息技术的强力驱动下, 数字人技术实现了跨越式发展。其形象精细度已达到以假乱真的水平, 智能化程度更是突飞猛进。如今, 从虚拟主播到数字员工, 从智能客服到 AI 助手, 各类数字人已深度渗透至多个不同行业领域, 在传媒、教育、医疗、金融等场景中扮演着日益重要的角色, 正加速融入人们的日常生活, 推动着人机交互方式的革新与升级。

1) 传媒领域中的应用。在传媒领域中, 基于完全虚拟的数字人的部分应用如下: 2024 年, 山东一千河数字科技开发的“海蓝手语数字人”系统入选中国正能量传播网络 AI 精品案例, 利用生成式 AI 技术为听障群体提供新闻资讯服务, 该系统通过 13000 多个词汇的手语库, 实现文本到手语的自动转化, 已在山东和宁夏广播电视台应用^[55]; 中国政府网推出 2 位 AI 集合多模态模型的数字人作为政策讲解员, 解读政府工作报告, 实现文字到视频的自动生成^[56]; 青岛互联文化传媒通过数字人名片, 生成动态形象, 用于个人和企业信息推介, 提升了社交分享率和信息承载量; 搜狐科技表示广州虚拟动力推出的 AIGC 3D 数字人视频创作方案, 帮助广电媒体提升制播效率和智能化水平, 该方案支持快速生成数字人播报视频, 适用于新闻、综艺等节目^[57], 如图 9 所示; 微软小冰作为社交网络虚拟伴侣, 展示了 AI 数字人在社交互动中的潜力^[58]; 粤传媒制作的《五羊传说》和《南粤水韵, 前年龙行》短片, 展示了 AI 技术在文化传播中的创新应用^[59]。



图9 3D 数字人进行新闻播报

基于真实人物的数字人部分应用包含: 2025 年, 两会期间, 全国人大代表张敏的专属数字人通过对接大模型(如 DeepSeek), 提供智慧对话服务, 如图 10 所示, 助力履职能力提升, 该数字人依托数字孪生和深度神经网络技术, 实现高度逼真的形象、音色和智能交互^[60]。从上述应用中可以发现数字人技术在传媒领域的应用正在重塑内容生产模式和传播形态, 为行业带来了效率提升和创意革新的双重机遇。未来, 随着技术的不断进步, 数字人将在更多场景中发挥重要作用, 推动传媒行业向智能化、个性化和多元化方向发展。

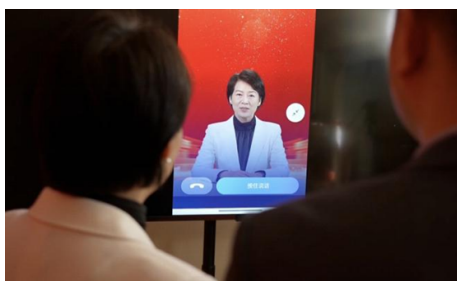


图10 基于真实人物的数字化复制品提供对话服务

2) 教育领域的应用。在教育领域, 数字人技术的应用正逐步改变传统的教学模式, 为学生和教师提供更加智能化、个性化和高效的学习与教学体验。中阿卫视使用数字人主播“毕诺-D”进行阿拉伯语新闻播报, 如图 11 所示, 将制作时间从数小时缩短至几十分钟, 显著提升了远程教育的效率^[61]; 部分高校利用数字人技术制作校园文化宣传视频, 通过数字人形象展示



图11 春节期间数字人播报新闻

学校的历史和特色, 吸引了更多学生和家长的关注^[62]; 风平智能的数字人可以为教育培训提供更加个性化的学习和培训体验, 同时还可以提高工作效率^[63]。

除了上述完全虚拟的数字人应用外, 世优科技的 AI 数字人系统帮助教师创建数字人分身, 快速生成教学视频, 将原本需要团队协作的内容制作流程简化为单人操作, 大幅提升了效率^[64]。任福继研究团队通过高精度数字人模型、实时语音识别、声音克隆、大语言模型、模型安全性检测和文本转语音模型, 构建了一个较为主流的实时交流的 AI 数字人, 如图 12 所示。其中高精度数字人利用动漫模型配合动作调整替代, 语音识别可采用科大讯飞或微信 AI 接口, 声音克隆采用 MockingBird 或 so-vits-svc 等模型, 大语言模型可采用 delta-tuning 对模型进行高效微调的方法来快速生成个性化对话模型, 采用量化的方式对模型运行的配置要求放宽, 采用语言安全监测方法来保证模型的安全性。数字人在教育领域的应用涵盖了从个性化学习到远程教育、从虚拟实验到情感陪伴的多个场景。这些应用不仅提升了教学效率和学习体验, 还推动了教育资源的公平分配和教育模式的创新。随着技术的进一步发展, 数字人将在教育领域发挥更加重要的作用。



图12 任福继形象的高精度 AI 数字人形象

数字人技术的发展已形成 2 大主要方向, 完全虚拟构建的数字人和实体映射数字人。无论是纯虚拟的数字人还是实体映射的数字人, 其技术演进和应用深化都离不开数字孪生场景的关键支撑。数字孪生通过构建物理世界的数字化镜像, 不仅为实体映射数字人提供了真实世界的运行基础, 也为完全数字人创造了与现实交互的接口环境。这种虚实融合的技术生态, 正在推动数字人向更智能、更协同的方向发展。

3 “场”:物与人融合构建的元宇宙生态

3.1 元宇宙的理论起源

元宇宙(metaverse)这一概念最早由美国作家尼尔·斯蒂芬森于1992年提出^[65],描述了一种与现实世界平行存在的虚拟世界,人们可在其中拥有独立身份并进行社会交往、经济交易、文化活动等。随着电影《Ready Player One》(《头号玩家》)的热播,元宇宙概念开始引起广泛关注^[66]。2021年,Facebook更名为Meta,标志着元宇宙成为全球科技领域的热门话题,许多科技公司纷纷加大在元宇宙技术方面的投入。元宇宙被认为是一种虚拟与现实无缝链接、深度融合的数字世界^[67]。尽管目前尚无统一的权威定义,但在现有研究中普遍认为元宇宙是通过科技手段创造的、与现实世界进行映射交互的虚拟空间,并在此基础上形成具有新型社会经济体系的数字生活环境^[68]。

从学术角度看,元宇宙研究体系通常被概括为6个主要方面,即区块链技术、交互技术、电子游戏技术、人工情感技术、网络及运算技术、物联网技术,这6大技术的英文首字母组成“BIGANT”,被称作元宇宙的“大蚂蚁”。Ren等^[69-70]在此基础上进一步完善,加入了人工情感技术,并提出了“A BIGANT”的元宇宙研究体系,如图13所示。

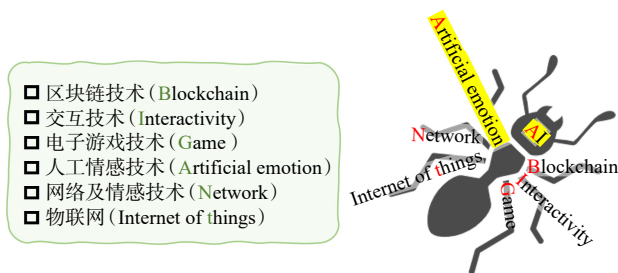


图13 元宇宙核心技术体系“A BIGANT”

元宇宙作为一种新兴的概念和技术领域,其关键特征主要包括以下4个方面。

1) 沉浸式体验:在元宇宙的构建中,高保真的虚拟现实技术带来了沉浸式体验,使用户不仅体验到视觉的逼真性,而且通过多感官的交互,如听觉和触觉,实现了超越现实的感官融合。技术进步将持续优化沉浸体验,模糊虚拟与现实的感官界限。

2) 虚实交互融合:元宇宙不仅是一个虚拟世界,更是与现实世界的无缝对接。通过物联网、5G、区块

链等技术,现实世界中的数据和信息可以实时地流入虚拟世界,反之亦然。例如,用户在现实世界中获取的信息、数据和行为,可以即时反映在虚拟空间中,而虚拟空间中的变化也可以对现实世界产生反馈。这样的虚实交互使得现实与虚拟之间形成了一个动态的、相互作用的闭环系统,为用户提供更加丰富的跨时空体验。

3) 社会性:元宇宙中的虚拟空间不仅是个人存在的场所,还是一个活跃的社交网络。用户在此可以进行多样化的社交活动,如虚拟聚会、文化交流等,形成了一个复杂的虚拟社会网络。随着AI和大数据技术的进步,用户的个性化社交圈得以构建,同时,社交网络的扩展超越了现实世界的地理、文化和语言限制,创造了全球性的交流空间。

4) 经济系统:元宇宙的经济体系以虚拟资产的创造、流通和交易为基础,构建了一个完整的虚拟经济生态系统。用户不仅能进行虚拟物品的买卖,还能创作并分享自己的数字资产。区块链技术的应用确保了虚拟资产的安全性、不可篡改性,促进了数字货币和非同质化代币(non-fungible token, NFT)等新兴数字资产的流通,推动了虚拟经济的全球发展和与现实经济的深度融合。

3.2 国家重大需求

全球各国均对元宇宙发展予以高度重视,并相继出台相关政策进行引导。从政策维度而言,中国已针对元宇宙建设颁布了一系列具有指导意义的政策,旨在推动其实现健康且有序的发展。2022年10月,五部委联合印发《虚拟现实与行业应用融合发展行动计划(2022—2026年)》,据预测,到2026年,中国虚拟现实产业的总体规模(涵盖相关硬件、软件、应用等)将超3500亿元人民币,虚拟现实终端累计销量将超过2500万台。自2023年起,国内元宇宙政策的出台持续保持2022年的快速增长态势,国家级以及地方层面的产业扶持政策与项目不断落地实施,为元宇宙产业的未来发展奠定了坚实根基,营造了优良环境。当前,中国已有上海、武汉、合肥、无锡、杭州、南昌、厦门等近40个省市市政府先后发布元宇宙建设规划,提出重点发展元宇宙相关产业,元宇宙成为各界瞩目的焦点,新模式新业态层出不穷。各行业对元宇宙开发平台与创作工作的持续完善为教育元宇宙的达成提

供了关键支撑。

国外政府已开始纷纷布局元宇宙。韩国政府明确目标,打造韩国成为全球5大XR经济国家,通过政府补贴、与企业合作的方式,推动元宇宙平台的开发和建设;印度在增强现实和虚拟现实领域的投资额达到1350亿印度卢比,每年增长38%;菲律宾信息和通信技术部鼓励创业公司加速数字化转型,为后续发展元宇宙奠定基础;印度尼西亚将数字战略定为未来印度尼西亚发展的重要战略计划,并将借G20峰会的时机,推出元宇宙平台Jagat;越南政府努力扩展元宇宙基础建设,鼓励越南的大型科技公司及初创公司向虚拟元宇宙大力投资,且为元宇宙的强大信息流做准备。同时,通过补充、修订及颁布大量法律工具的全面立法改革过程体现了对虚拟技术发展的坚定承诺。基于此,各国政策的颁布为教育的数字化转型的实现提供了政策支持。

尽管各国政策走向各有侧重,却无不致力于元宇宙的有序化布局,旨在促进其规范化的进步。可以预见,大力发展元宇宙将成为未来数字经济的重要趋势。

3.3 数字孪生技术在元宇宙场景构建中的关键作用

数字孪生与元宇宙是近年来备受瞩目的2个概念,二者既存在诸多相似之处,亦有不少差异。数字孪生是一项融合实际物理对象或系统数字模型与实时数据的技术,旨在达成对实际对象或系统的实时监测、预测及优化。元宇宙则是以虚拟现实、增强现实等技术为基石构建的虚拟数字世界,由多个数字孪生模型构成,这些模型对现实世界的物理、空间、人类活动及社会互动予以模拟,能使用户于虚拟空间中获得更为真实的体验,实现与其他用户的互动协作,同时也为商业活动创造了更多契机。

从技术演进维度审视,元宇宙的构建可划分为具有递进关系的3个技术层级,即数字孪生层、虚实交互层与元宇宙生态层。其中,数字孪生层作为基础阶

段,在元宇宙架构中占据着极为关键的底层支撑地位,其发挥的关键作用如表4所示。然而,要实现可交互的数字孪生和全面的元宇宙体验,动态场景的数字孪生成为技术实现的核心突破。动态场景的数字孪生不仅仅是物理实体的数字映射,而是通过实时感知、建模和反馈现实世界的变化,推动虚拟环境与真实世界的深度融合。在此过程中,虚拟世界中的场景和环境能够实时响应并调整以契合现实世界的动态变化,使得虚拟空间具备“自适应”能力。为实现上述目标,动态场景的数字孪生依赖于多种技术和表示方法,如动态网格、4D NeRF以及高斯散射模型等,通过实现时间与空间的几何一致性,确保虚拟与现实的精准映射^[7]。这些技术为动态环境的生成提供了强有力的支持,进而实现了虚拟空间与物理世界的无缝连接。通过这些创新,用户不仅能够在虚拟空间中享受到更加沉浸的体验,还能够在虚拟环境中的场景和其他用户进行流畅、实时的互动。由此,动态场景的数字孪生技术为元宇宙中的交互性、真实性和沉浸感提供了强有力的技术保障。

3.4 数字人与元宇宙交互融合技术分析

随着元宇宙技术架构的迭代更新,数字人作为连接现实世界和虚拟空间的关键媒介,在提升虚拟角色交互性与沉浸感的同时,也为元宇宙社交场景中的情感计算与反馈机制提供了新的技术实现路径。随着人工智能、虚拟现实以及增强现实技术的持续发展,数字人的应用场景也逐渐从简单的虚拟角色扩展至更多元化的功能,如虚拟代言人、品牌推广、虚拟顾问等。数字人在元宇宙中的应用大大提升了用户的参与感和沉浸感,成为构建更加真实和互动的虚拟世界的关键要素(表5)。

3.5 “物-人-场”三元交互的数据流框架

在综合分析“物”“人”“场”3个要素的基础上,提出了“物-人-场”三元交互模型。该模型由3个核心

表 4 数字孪生技术在元宇宙场景构建中的关键作用

主要作用	具体表现
虚拟世界构建	通过高精度3D建模和数据映射,实现现实世界的数字化复刻
实时数据同步	借助物联网和传感器技术,实现物理世界与虚拟世界的交互
智能仿真	模拟物理环境、用户行为及系统响应,优化决策和运营
沉浸式交互	借助VR/AR/MR技术提升用户沉浸感,增强虚拟体验

表 5 数字人在元宇宙场景构建中的关键作用

主要作用	具体表现
虚拟角色交互	通过数字人实现用户与虚拟世界的实时互动与交流,提供个性化互动体验
沉浸式体验增强	数字人通过实时渲染和情感计算技术,增强用户的沉浸感与真实感
社交功能赋能	支持多人在线互动,数字人参与群体互动并模拟社交行为
情感计算与反馈	数字人能够通过情感分析与理解,适应用户情绪并进行回应

模块构成:物理实体、人(自然人和数字人),以及虚实融合环境,并以高效的数据交互网络作为连接各模块之间的纽带,构建起一个多模态协同、动态响应的交互系统框架,如图 14 所示。该模型中的交互网络主要涉及 3 类关键数据流路径。



图 14 “物-人-场”三元交互的数据流框架

1) “物”到“场”的数据流路径主要体现在物理实体对虚拟环境的实时映射能力。

以教育元宇宙为例,该路径的典型应用是将真实教室中的教学硬件设施(如智能黑板、投影设备、实验仪器、环境监控系统等)通过传感器与物联网技术进行数字化建模与状态采集。这些传感设备可实时采集温湿度、光照、噪声、设备使用状态等信息,生成的环境感知与运行数据首先由边缘计算节点进行预处理与筛选,随后传输至云端融合计算平台完成数据集成与场景建模。最终,这些数据将驱动教育元宇宙中数字教室的三维仿真渲染,使得虚拟空间能够精准同步现实教室的运行状态与教学环境,确保虚实之间在教学场景感知上的高度一致性与动态联动。

2) 从“人”到“场”的数据流路径聚焦于自然人、数字人与虚拟环境之间的智能交互过程。

在教育元宇宙场景中,该路径典型地表现为用户通过动作捕捉设备、语音识别系统、表情摄像头等多模态交互装置,与虚拟教学环境实现实时互动。系统可采集自然人的语音指令、肢体动作(如举手、转头)、面部情绪(如专注、疑惑等),同时也记录数字人在环境中的交互行为与语言响应。这些多模态数据首先在边缘 AI 节点上进行预处理,以降低时延,随后

被传输至 AI 交互引擎,融合自然语言理解、情感识别与行为建模等算法,对用户意图、认知状态与情绪变化进行智能分析。最终,系统在虚拟环境中驱动数字人角色(如虚拟教师)进行实时回应,包括知识讲解、语音反馈、表情表现或虚拟动作执行,从而构建起以自然人和数字人协同参与的高沉浸式交互课堂,显著提升教育元宇宙中的教学真实感与个性化水平。

3) 从“场”反馈至“物”与“人”的数据流体现了虚拟环境的响应能力与闭环控制特征。

在教育元宇宙中,虚拟教学环境不仅是交互的承载空间,还具备基于实时感知与动态分析的反馈调节功能。随着教学活动的进行,虚拟环境不断生成场景交互数据、学习行为分析结果及控制指令,例如学生的专注度评估结果、知识点掌握情况、交互频率变化等。这些数据经由云端决策模块综合分析后,分别反馈至现实中的物理教学设备与虚拟教学系统。在“物”端,系统可自动调节教室内的智能设备状态,如调节灯光亮度、调整教学终端的界面内容或启动实验仪器,营造适宜的学习环境;在“人”端,反馈数据则用于实时调控虚拟教师(数字人)的讲授策略与交互行为,例如,根据学生情绪状态调整语速和讲解方式,或根据群体认知状态生成针对性的补充讲解与提问策略,从而实现因材施教与个性化响应。通过这一数据反馈机制,教育元宇宙实现了虚实空间之间的协同闭环控制,显著增强了教学互动的主动性、适应性与智能化水平。

3.6 国内外典型元宇宙应用案例

随着“物-人-场”三元交互系统理论的逐步完善,虚实融合的元宇宙形态也正在由概念走向应用。可以说,元宇宙的核心特征之一正是“物理实体(物)—智能主体(自然人和数字人)—虚实环境(场)”三者在时空上的实时感知、认知、交互与反馈。而数字孪生技术的演进,正为这一融合提供了底层支撑与技术路

径。以下将从游戏、教育与商业3个领域出发,梳理国内外典型案例,深入探讨“物-人-场”融合在不同场景中推动的元宇宙实践路径。

游戏领域是当前元宇宙应用最为成熟的领域之一,凭借其天然的交互性与沉浸性,为元宇宙的场景搭建、身份建构与经济体系探索提供了理想试验场。国外典型代表如 Roblox 和 Fortnite,通过构建整合用户生成内容(user-generated content, UGC)、交易与社交于一体的开放式虚拟平台,为用户提供了高度自由的数字体验空间。这些平台不仅引入了虚拟演唱会、数字资产等功能,也在不断探索元宇宙经济体系的可行路径。国内方面,以《原神》《幻塔》等开放世界游戏为代表,通过引入 AI 智能非玩家角色(non-player character, NPC)、数字服饰定制与实时渲染等新兴技术,推动游戏从传统线性叙事向持续演化的沉浸式虚拟世界转型。此外,娱乐圈的大牌明星如贾斯汀·比伯、马什梅罗和特拉维斯·斯科特等,也纷纷投身于元宇宙娱乐的探索,推动沉浸式娱乐体验不断创新与发展。这些探索不仅展现了“物-人-场”融合下元宇宙实践的活力,也为其他行业提供了跨界参考路径。

在游戏领域率先实现元宇宙技术落地的基础上,教育场景成为元宇宙融合发展的重要延展方向之一。与游戏注重娱乐性与沉浸感不同,教育类元宇宙更强调教学功能性与学习效果的提升。其核心在于打破传统教育中时间、空间和资源的限制,通过虚拟现实、增强现实等技术提供更具互动性和参与感的学习体验。国外如 Engage VR 和 Meta 的 Horizon Classrooms 等平台,借助 VR/AR(虚拟现实/增强现实)技术构建可多人实时互动的虚拟课堂,支持跨地域协作教学和学科资源共享;同时,沉浸式语言学习、虚拟科学实验等应用场景已被广泛验证其教学有效性。具体来看,以欧洲高校为例,2024年9月,英国利兹大学率先开设“元宇宙戏剧”课程,学生通过佩戴 VR 设备参与虚拟排练与演出,实现了对传统舞台空间的突破;2025年2月,西班牙巴斯克大学推出基于虚拟现实的解剖学课程,使学生能够在 3D 虚拟空间中观察人体器官结构并完成模拟手术;此外,德国汉诺威地区多所高校也陆续上线沉浸式课程,覆盖医学、工程等多个专业领域,支持学生通过虚拟实验室完成复杂实验任务,极大地提升了教学的可操作性与

沉浸感。国内高校和科技企业也积极响应元宇宙发展潮流,推动其在教育领域的深度应用,涌现出一批兼具本土特色与技术创新的示范项目。例如,清华大学打造了“元宇宙校园”,实现了校园空间与教学设施的数字化复刻,支持虚拟实验室、数字导览等创新教学方式,广泛服务于新生引导及专业课程;南开大学则开设了国内首个本科层面的元宇宙选修课程,并联合企业共建虚拟课堂及数字人实验室,致力于培养虚拟直播等新兴技能;山东大学开发了虚拟校园导览系统,使新生能够通过 AR 技术实时体验校园建筑 and 历史文化场景;西北大学建设了陕西省首个元宇宙校园平台,支持虚拟会议、在线实验及数字文创产品发行,覆盖工程与医学等多个学科领域。

随着元宇宙的不断成熟,除了教育领域的广泛应用,商业领域的应用日益丰富多样,涵盖了零售、房地产、制造、金融等多个行业。国外方面,耐克以“数字替身+虚拟产品+Nikeland”重构品牌营销要素,借助数字孪生技术,用户不仅可以个性化定制数字形象,还能享受丰富的虚拟互动体验,同时实现虚拟试鞋和沉浸式购物,极大提升了品牌的数字化参与感和用户黏性;Matterport 利用数字孪生技术,为房地产经纪公司和门户网站打造虚拟 3D 看房体验,显著提升了线上用户的参与度和访客转化率,有效推动了房地产展示的数字化升级与效率提升;西门子的虚拟协同仿真元宇宙依托其先进的数字孪生技术,实现产业链上下游各环节在统一虚拟环境中的实时动态模型调整。该平台支持产品研发设计与生产线升级的高效协作,突破了时间与空间的限制,显著提升了协同效率,降低了运营成本;摩根大通成为华尔街首家进军元宇宙的银行,打造了基于区块链平台 Decentraland 的“Onyx 休息室”虚拟空间。在 Decentraland 中,用户可通过加密货币以 NFT 形式购买虚拟土地。该“休息室”由摩根大通旗下区块链部门 Onyx 负责建设,旨在为用户提供沉浸式的虚拟互动体验。国内则涌现出阿里巴巴虚拟购物空间和数字藏品经济、碧桂园智慧社区数字孪生管理、华为智能工厂数字化升级,以及建设银行在数字货币与虚拟资产交易上的创新实践。这些典型案例充分展示了“物-人-场”融合在商业元宇宙中的应用潜力,推动传统产业向数字经济转型升级。

3.7 教育元宇宙生态的应用案例

教育元宇宙作为新兴的教育模式,正借助数字技术的发展,逐步打破传统教育的时空限制,为学生和教育工作者提供了全新的沉浸式学习体验。本文以教育元宇宙为例,深入探讨元宇宙生态的构建和应用。在教育元宇宙的构建中,“物-人-场”三重交叉关系不仅使每个要素独立发挥作用,更通过相互协同使得教育元宇宙能够真正实现沉浸式、个性化、智能化的教育模式,从而突破传统教育的时空限制,带来前所未有的教育体验,如图 15 所示。其中,“物”指真实教育场景的数字孪生,“人”包括智能虚拟教师、虚拟学生以及真实学生的数字孪生映射,“场”则是虚实融合的教育环境,实现了跨地域、个性化、沉浸式的学习体验。以下将分别展开介绍这 3 大要素在教育元宇宙的交互关系。

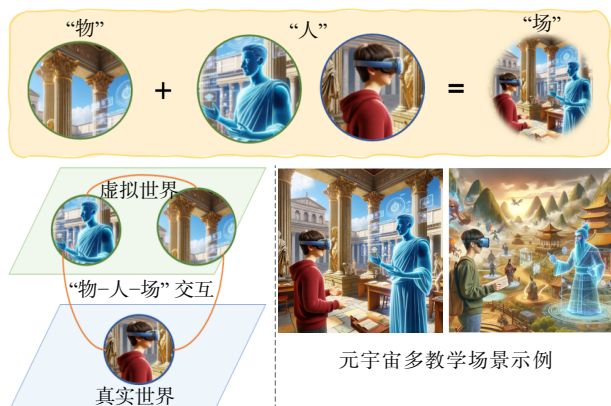


图 15 教育元宇宙的组成示例

1) 物与人:物理环境和数字人的交互。在教育元宇宙架构下,物品的数字映射与虚拟人物的交互构成了教育活动的核心要素。一方面,虚拟教师与现实环境之间的互动表现为虚拟教师能在数字教室复制品中“实施”教学活动,并通过智能感知技术的应用对教学素材进行动态调整。例如,在物理学课堂上,虚拟数字教师可以根据实验设备的状态(如数字孪生的化学试剂瓶、机械装置等)灵活调整实验步骤和内容,从而使实验结果能够更加精准地反映学生的操作情况。另一方面,学生的虚拟化身(数字孪生)能够直接与数字孪生实验设备进行交互。以虚拟化学实验为例,学生通过数字孪生体在虚拟实验室中进行实验操作,实验设备(如试管、烧杯等)会实时反馈数据。虚

拟教师则根据学生的操作结果提供即时反馈与指导。通过这种物与人之间的交互,学生能够身临其境地进行实践操作,而教师则可以根据学生的虚拟行为和学习进度动态调整教学内容,从而显著提升教育的个性化水平。

此外,为了实现更加真实与沉浸的教学体验,教育元宇宙中的交互不应仅局限于虚拟空间,还需进一步拓展至现实用户与物理场景的融合交互。例如,通过部署多维度传感器(如动作捕捉、触觉传感、温度感应等),可以实时感知学生在现实环境中的动作状态,并将这些行为同步映射至虚拟空间;同时,引入具备物理反馈能力的交互装置(如力反馈手套、震动回馈设备)也可将虚拟实验中的关键操作反馈至学生的真实感官系统,形成虚拟与现实间的闭环。进一步,教育活动中的“人-物”交互将从传统的视觉/听觉层面延伸至触觉与体感层面,从而有效提升学习的沉浸感、交互性与个性化响应能力。

2) 人与场:虚拟人与虚拟场景的融合与互动。在教育元宇宙的构建中,“场”不再局限于静态的教育场所,而演变为一个集动态交互与智能技术于一体的动态环境。数字化身与虚拟场景的深度整合,显著提升了学习体验的沉浸性与个性化特征。一方面,在虚拟教育环境中,学生得以通过其数字化身便捷地参与跨国界的虚拟课堂,并与来自世界各地的同龄人进行实时交流。虚拟课堂能够根据学科特性灵活调整教学场景,如在历史课程中模拟古罗马时期的教学场景,在物理课程中则切换至现代实验室。这种灵活直观的学习模式使学生得以依据个人兴趣自主选择学习环境。此外,数字人(如智能教学助手)为学生量身定制个性化的学习路径,并在特定学习场景中扮演指导者、历史人物等角色,提供全面的学习引导和支持。另一方面,“场”为学生营造出沉浸式的互动体验。以医学教育为例,虚拟场景能够精确模拟人体解剖实验室,学生虚拟化身可自由探索这一环境,虚拟医生或学长在场景中提供专业指导与即时解答。学生通过虚拟现实技术与实验环境的直接互动,在感官上体验到沉浸式的学习感受。这种人与环境的交互模式,不仅激发了学生的主动学习热情,而且场景的多样性和互动性显著增强了学习的趣味性和实效性,同时有效提升了学生的实践操作能力和跨学科思维能力。

进一步地,虚拟场景的智能响应能力使其不仅作为教学的背景环境存在,更逐步演化为具备感知与反馈能力的主动教学参与体。当系统通过多模态传感设备(如动作捕捉、眼动追踪、语音识别等)感知到学生在现实物理空间中的行为状态,例如,在特定区域的停留时间、视线聚焦点或语音交流频率时,虚拟场景中的内容会据此实时生成、扩展或调整,以契合其当前的学习阶段与行为意图。这种“真实人—虚拟场”的感知联动模式,不仅拓展了传统教育场景的感官边界,更在实践操作与情境理解层面激发了学生的主动参与意识。通过行为感知与场景动态适配的耦合机制,教育过程更具响应性、个性化与沉浸感,从而有效提升学生的操作能力、情境理解力与跨学科整合素养。

3) 物与场:物理环境与虚拟教育场景的融合。在教育元宇宙的构建中,物理环境的数字化映射与虚拟环境的融合构筑了一种新型的虚实结合的教育互动空间。一方面,虚拟实验室与实际设备的交互为学生提供了一个全新的实验平台。学生得以在虚拟环境中操控虚拟设备,进行物理实验的模拟操作,如化学实验中的反应流程、物理实验中的力学现象等,均能通过数字孪生技术在虚拟环境中得以呈现。虚拟实验室的数据来源不仅限于虚拟设备,还与真实实验室的数据实现交互,为学生提供了更为精确的模拟体验。另一方面,实体教室与虚拟课堂的协同作用扩展了教育活动的边界。在某些课程中,教育元宇宙中的虚拟课堂与物理教室可实现实时互动。学生可在物理教室参与远程虚拟课堂,与远端教师及全球学生展开合作。物理教室通过感应器(如温度、湿度、空气质量等)向虚拟课堂传输实时环境数据,而虚拟课堂则依据这些数据调整教学内容。这种物理与虚拟环境的深度融合,不仅超越了物理与虚拟空间的简单叠加,更创造出一种灵活、智能、互动的教育新模式。

4 “物-人-场”交互视域下数字孪生技术的挑战与趋势

4.1 “物-人-场”交互面临的研究挑战

数字孪生技术经历了从传统的“物”的单一对象仿真,到AI驱动的“人”的智能交互,再到虚拟融合的“场”生态建设的发展阶段。尽管各发展阶段都取得

了显著进展,但如图16所示,在“物-人-场”融合的新范式下,仍面临诸多共性技术挑战。

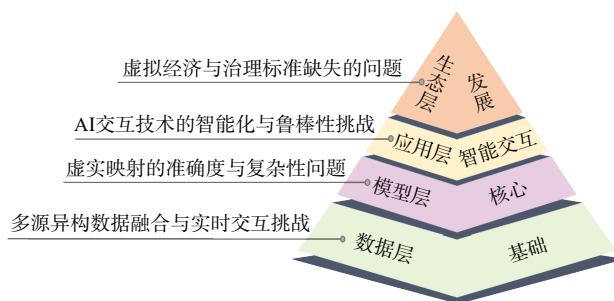


图16 “物-人-场”交互面临的研究挑战

1) 多源异构数据融合与实时交互挑战:在“物-人-场”深度融合的教育元宇宙与数字孪生系统中,实现高效、低延迟的多源异构数据融合是系统智能化的基础。然而,数字孪生技术在实际应用中面临着复杂的数据收集与整合挑战。首先,由于数据来源广泛,涵盖结构化表格、非结构化音视频、半结构化日志等多种形式,数据格式缺乏一致性,不同模态数据难以直接拼接与对齐,显著增加了统一建模的难度。其次,数据质量参差不齐问题普遍存在,诸如传感器误差、缺失值、噪声干扰等均可能对融合结果的准确性和稳定性产生负面影响,进而削弱系统在关键任务中的表现。此外,数据在采集、传输和存储过程中面临安全与隐私隐患,如身份信息泄露、恶意篡改或非法利用等风险,严重制约了数字孪生系统的可扩展性与可信度。

为此,亟需构建具备鲁棒性的多模态数据处理架构,引入数据清洗与标准化机制,提升数据质量的一致性与可靠性;同时,加强数据生命周期管理,借助加密存储、访问控制与联邦学习等技术,保障数据在使用过程中的隐私与安全。只有解决上述基础性问题,才能为实现泛在感知、智能决策与人机协同交互提供可信的数据支撑。

2) 虚实映射的精准度与复杂度问题:高精度虚拟建模与实时仿真是数字孪生系统实现“真实世界映射—虚拟世界预测—现实控制反馈”闭环的核心能力。然而,在“物-人-场”融合背景下,系统对时空建模精度、交互响应速度以及可扩展性的要求显著提升,造成虚实映射准确性与系统复杂度之间的矛盾愈发突出。数字孪生的本质在于将物理世界中的设备、

个体和环境等要素高保真地投射至虚拟空间,并实现实时状态同步、未来状态预测与反馈控制优化。然而,这一过程往往伴随着极高的计算开销与系统负担,尤其在涉及复杂人机交互与多主体协同场景时更为明显。

具体而言,当前主要面临以下3个挑战:(1)高保真建模需要大量多源传感数据支持,并依赖精细网格建模、复杂的纹理贴图与物理仿真逻辑,导致模型规模庞大、部署与运算效率大幅下降,特别在工业装备、复杂设施等场景中尤为突出;(2)实时渲染与交互处理对计算资源的依赖日益增强,尤其在涉及数字人建模(如面部表情肌肉驱动、情绪驱动行为建模)与自然语言处理等交互任务时,需依赖高性能GPU(图形处理单元)、边缘AI计算节点等硬件支持,中小型平台或教育场景难以承载;(3)多用户并发访问虚拟环境(如元宇宙教室、虚拟工厂、沉浸式会议)将带来巨大的系统负载,可能引发交互卡顿、画面撕裂、数据同步延迟等问题,严重影响整体体验一致性与沉浸感;在一些应用中,为了在有限算力条件下实现近似实时的仿真,往往需要在仿真精度与系统延迟之间进行权衡,这进一步加剧了建模精度与系统性能之间的张力。因此,未来数字孪生系统的建设亟需探索轻量级建模方法、分布式渲染技术、多层次算力协同机制以及高效的数据压缩与同步算法,才能在保持虚实映射精度的同时,实现系统运行的高效性、稳定性与可扩展性。

3) AI交互技术的智能化与鲁棒性挑战:在“物-人-场”交互融合体系中,人工智能交互技术作为数字人驱动、人机融合,以及环境适配的核心引擎,正日益发挥关键作用。然而,当前AI技术在语义理解、情境感知和人机协同等核心环节的智能化水平与系统鲁棒性仍存在显著不足,已成为制约复杂交互系统稳定运行与广泛落地的主要瓶颈之一。

首先,现有AI交互模型在复杂场景下的鲁棒性表现仍显脆弱。特别是在开放式、多模态的人机交互环境中,用户输入往往具有高度非结构化、不确定性与语义歧义,容易导致AI模型在语义解析、指令理解等环节出错。例如,在自然语言处理任务中,模型对口语化表达、情绪化语言和上下文跳跃的理解能力不足,导致指令响应不准确;在图像识别与行为建模中,

模型易受光照、姿态遮挡或场景干扰等因素影响,识别准确率显著下降。其次,AI交互模型的泛化能力有限,难以应对跨场景迁移与新输入模式的挑战。在多数应用中,AI模型往往依赖训练集中的“已知模式”进行推理与响应,但在教育元宇宙、智能制造或应急指挥等典型场景中,用户交互行为高度多样且动态变化,若缺乏足够的领域适应性与增量学习能力,模型在遇到“未知指令”或“新场景”时容易表现出错误响应、无效反馈甚至系统崩溃。这一问题在自动驾驶中的意外场景判断、情感识别中的跨文化表达理解等领域尤为明显。最后,当前AI交互系统尚缺乏稳定高效的多模态协同机制,语音、图像、动作等感知通道之间存在数据融合与时序同步障碍,影响整体交互流畅性与语义一致性。同时,AI模型仍面临算力消耗高、部署灵活性差、训练数据需求庞大等现实限制。因此,提升AI交互模型的鲁棒性与泛化能力,需要从多层面协同突破:包括引入知识增强推理机制、实现少样本学习与自适应迁移、构建多模态信息融合架构、优化模型压缩与部署策略等。未来,唯有在保持高精度交互的同时兼顾可解释性与稳定性,AI系统才能真正支撑“物-人-场”融合环境中的智能化、可持续的人机协作。

4) 虚拟经济与治理标准缺失的问题:随着元宇宙生态体系的不断扩展,虚拟经济正逐步成为其核心支柱之一。数字孪生与数字人技术的深度融合,不仅加速了虚拟资产(如数字地块、虚拟商品)与虚拟身份的生成与流通,也推动了虚拟社会中交易活动、身份体系与价值网络的初步构建。然而,目前该领域在权属确权、交易规范、跨平台互操作性及资产安全保障等方面仍存在明显的制度空白与技术滞后,已成为制约元宇宙可持续发展的关键治理瓶颈。一方面,虚拟资产的所有权、使用权和收益权界定尚不清晰,缺乏具有法律效力的确权机制和数字原生认证标准,容易引发资产归属争议与产权纠纷。不同平台之间的资产转移和身份互认机制尚未打通,导致用户资源与价值受限于平台孤岛效应。另一方面,虚拟经济活动中存在高频交易、去中心化结算与高价值资产交换的特性,但其配套的风险控制机制、安全协议与监管框架尚未成熟,面临身份伪造、洗钱、诈骗、数字资产丢失等多重风险。此

外,当前尚缺乏国际统一的虚拟经济与虚拟身份治理标准,不同国家和地区在数据主权、身份合规、金融监管方面存在巨大差异,给全球化虚拟经济的治理与协同监管带来严峻挑战。因此,未来需要从技术与制度2个维度同步推进:在技术层面,应构建可信身份认证机制、可验证数字资产标准以及跨平台的资产互通协议;在制度层面,应加快构建虚拟经济相关法律体系与治理机制,推动建立面向元宇宙环境的国际通用标准与多边监管合作框架,从而实现虚拟经济的合规发展与风险可控。

4.2 “物-人-场”交互视域下未来研究方向

基于上述挑战,为推进数字孪生技术正从传统的静态仿真模型向智能化、自主进化的虚实融合系统演进,未来研究可以聚焦在以下3个方面。

1) 构建智能化且具备多层次架构的数据融合框架。针对数据孪生技术在数据收集阶段面临的数据格式缺乏统一标准、数据质量良莠不齐等难题,构建数据格式标准化协议,诸如基于工业互联网数据格式标准(如ISO 23247)以及知识图谱驱动的数据映射技术,达成不同类别数据的自动校准,削减数据整合的复杂度,促使跨域、跨模态数据实现无缝融合;针对数据存在安全风险的状况,运用区块链、联邦学习以及多方安全计算等技术手段,提升数据的安全性与可信度。为实现多源数据的更优融合,可从局部传感器至全局数字孪生模型设计多层次的数据融合架构。

2) 研究AIGC(生成式人工智能)赋能的智能虚实映射与虚拟原生进化。AIGC作为推动虚拟现实发展的重要力量,不仅预示着虚拟世界发展新纪元的到来,而且能够通过精确再现现实,甚至创造出现实世界所不具备的虚拟场景,进而孕育出“虚拟原生”的新型模式。面对虚实映射的精度与复杂性挑战,AIGC能够自主生成新的数字孪生,巧妙地克服了相关难题;针对现有数字孪生在泛化能力上的不足,AIGC凭借其内容生成能力超越了时空界限,激发了创新思维的涌现,创造了全新的虚拟场景;同时,在“物-人-场”交互视域下,AIGC能够构建出有助于深度理解的生动场景,从而增强用户的沉浸体验与互动热情。

3) 构建新的虚拟经济与智能化治理体系。在数字孪生技术与AIGC驱动的数字人集成背景下,预计虚拟经济领域将构建起更为完善的产业架构。一方

面,区块链与智能合约的应用有望解决虚拟资产的所有权界定及交易过程的透明度问题;另一方面,AIGC技术对虚拟资产交易数据的分析有助于自动揭示交互环节中的欺诈隐患。此外,通过在元宇宙中为数字人构建身份体系,尤其是借助AIGC生成个性化的数字身份,结合分布式存储技术保障隐私安全,将有助于构建一个统一、可信、可控的身份管理框架,进而推动跨平台协作与经济流动的深化。

综上,AIGC技术的融合正促使数字孪生技术从基于数据的静态模拟系统,向智能化融合的虚实一体的生态系统发展。未来,随着大模型、联邦学习、边缘计算等技术的持续突破,AIGC+数字孪生将具备更强的自主学习、决策优化和创造新环境的能力,在工业、医疗、教育、智慧城市等领域释放更大的价值。这一融合将不仅改变人们的生产、生活方式,也将催生全新的智能社会形态,使虚拟与现实真正融合,构建出更加智能、高效、沉浸式的数字未来。

5 结论

数字孪生技术作为工业制造、人工智能和虚拟现实等多个领域高度关注的前沿技术,正不断推动虚拟与现实世界之间的深入融合。从早期聚焦于工业设备和生产流程的传统“物”的数字孪生,到以AI驱动的数字人为核心的“人”的数字孪生,再到数字孪生与数字人融合构建的虚拟“场”(元宇宙生态)，“物-人-场”的交互视域清晰地体现了数字孪生技术发展的完整路径与演变趋势。围绕这一内在的发展逻辑,系统地综述了数字孪生技术在不同发展阶段的研究进展与技术细节,并总结了“物-人-场”交互视域下数字孪生技术所面临的共性技术挑战,从技术发展趋势视角进行预测。

参考文献(References)

- [1] Grieves M. Digital twin: Manufacturing excellence through virtual factory replication[R/OL]. [2020-12-01]. <https://www.3ds.com/fileadmin/PRODUCTS-SERVICES/DELMIA/PDF/Whitepaper/DELMIA-APRISO-Digital-Twin-Whitepaper.pdf>.
- [2] Singh M, Fuenmayor E, Hinchy E, et al. Digital twin: Origin to future[J]. Applied System Innovation, 2021, 4(2): 36.

- [3] Gelernter D. Mirror worlds: Or the day software puts the universe in a shoebox—how it will happen and what it will mean[M]. Oxford: Oxford University Press, 1993.
- [4] Grieves M, Vickers J. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems[M]// Transdisciplinary Perspectives on Complex Systems. Cham: Springer International Publishing, 2016: 85–113.
- [5] Kritzing W, Karner M, Traar G, et al. Digital twin in manufacturing: A categorical literature review and classification[J]. IFAC–PapersOnLine, 2018, 51(11): 1016–1022.
- [6] Tao F, Xiao B, Qi Q L, et al. Digital twin modeling[J]. Journal of Manufacturing Systems, 2022, 64: 372–389.
- [7] Sun Z Y, Zhang R T, Zhu X M. The progress and trend of digital twin research over the last 20 years: A bibliometrics–based visualization analysis[J]. Journal of Manufacturing Systems, 2024, 74: 1–15.
- [8] Shao G D, Helu M. Framework for a digital twin in manufacturing: Scope and requirements[J]. Manufacturing Letters, 2020, 24: 105–107.
- [9] Liu Y, Feng J, Lu J M, et al. A review of digital twin capabilities, technologies, and applications based on the maturity model[J]. Advanced Engineering Informatics, 2024, 62: 102592.
- [10] 清华大学互联网产业研究院. 数字孪生技术应用白皮书[R/OL]. [2023–03–24]. https://eng.pbcsf.tsinghua.edu.cn/_local/F/03/C6/D0A8223D74570BB2C192271AE97_1FE2A164_23BCA3.pdf.
- [11] 艾瑞咨询研究院. 中国数字孪生行业研究报告[R]. 北京: 艾瑞咨询, 2023.
- [12] Yi Y, Zhang A Q, Liu X J, et al. Digital twin–driven assembly accuracy prediction method for high performance precision assembly of complex products[J]. Advanced Engineering Informatics, 2024, 61: 102495.
- [13] Kazhdan M, Bolitho M, Hoppe H. Poisson surface reconstruction[C]//Proceedings of the fourth Eurographics Symposium on Geometry Processing. Goslar: Eurographics Association, 2006, 7(4).
- [14] Poole B, Jain A, Barron J T, et al. Dreamfusion: Text–to–3D using 2D diffusion[C]//The Eleventh International Conference on Learning Representations. Kigali: Openreview, 2023.
- [15] Tao F, Qi Q L, Wang L H, et al. Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison[J]. Engineering, 2019, 5(4): 653–661.
- [16] Meng W, Yang Y L, Zang J Y, et al. DTUAV: A novel cloud–based digital twin system for unmanned aerial vehicles[J]. Simulation, 2023, 99(1): 69–87.
- [17] Pan Y D, Braun A, Borrmann A, et al. 3D deep–learning–enhanced void–growing approach in creating geometric digital twins of buildings[J]. Proceedings of the Institution of Civil Engineers–Smart Infrastructure and Construction, 2023, 176(1): 24–40.
- [18] Grieves M W. Product lifecycle management: The new paradigm for enterprises[J]. International Journal of Product Development, 2005, 2(1/2): 71.
- [19] Kapteyn M G, Pretorius J V R, Willcox K E. A probabilistic graphical model foundation for enabling predictive digital twins at scale[J]. Nature Computational Science, 2021, 1(5): 337–347.
- [20] Zienkiewicz O C, Taylor R L, Zhu J Z. Adaptive finite element refinement[M]//The Finite Element Method: Its Basis and Fundamentals. Amsterdam: Elsevier, 2013: 545–572.
- [21] Wang Y, Sun Y B, Liu Z W, et al. Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics, 2019, 38(5): 1–12.
- [22] Ganin Y, Bartunov S, Li Y, et al. Computer–aided design as language[J]. Advances in Neural Information Processing Systems, 2021, 34: 5885–5897.
- [23] Kadry K, Gupta S, Nezami F R, et al. Probing the limits and capabilities of diffusion models for the anatomic editing of digital twins[J]. NPJ Digital Medicine, 2024, 7: 354.
- [24] Xu H W, Shao Y L, Chen J F, et al. Semi–automatic geographic information system framework for creating photo–realistic digital twin cities to support autonomous driving research[J]. Transportation Research Record: Journal of the Transportation Research Board, 2024, 2678(6): 1068–1084.
- [25] Sun Q, Das S, Wang K, et al. Toward a Spatial metaverse: Building immersive virtual experiences with georeferenced digital twin and game engine[M]//Big Data. Boca Raton: CRC Press, 2025: 334–360.
- [26] Müller T, Rousselle F, Novák J, et al. Real–time neural radiance caching for path tracing[J]. ACM Transactions on Graphics, 2021, 40(4): 1–16.
- [27] Müller T, Evans A, Schied C, et al. Instant neural graphics primitives with a multiresolution hash encoding[J]. ACM Transactions on Graphics, 2022, 41(4): 1–15.
- [28] Tewari A, Thies J, Mildenhall B, et al. Advances in neural rendering[J]. Computer Graphics Forum, 2022, 41(2): 703–735.
- [29] Wu S A, Basu S, Brodermann T, et al. PBR–NeRF: Inverse rendering with physics–based neural fields[C]//Online: Computer Vision and Pattern Recognition, 2024.
- [30] Bieron J, Tong X, Peers P. Single image neural material relighting[C]//Proceedings of Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Proceedings. New York: ACM, 2023: 1–11.
- [31] Yang J, Xiao H, Teng W, et al. Light Sampling field and brdf representation for physically–based neural rendering[C]//The Eleventh International Conference on Learning Representations. Kigali: Openreview, 2023.
- [32] Qin Y, Liu H Y, Wang Y, et al. Inverse physics–informed neural networks for digital twin–based bearing fault diagnosis under imbalanced samples[J]. Knowledge–Based Systems, 2024, 292: 111641.
- [33] Eyring A, Hoyt N, Tenny J, et al. Analysis of a closed–loop digital twin using discrete event simulation[J]. The International Journal of Advanced Manufacturing Technology,

- 2022, 123(1): 245–258.
- [34] Zhang H, Qi Q L, Tao F. A multi-scale modeling method for digital twin shop-floor[J]. *Journal of Manufacturing Systems*, 2022, 62: 417–428.
- [35] Es-haghi M S, Animescu C, Rabczuk T. Methods for enabling real-time analysis in digital twins: A literature review[J]. *Computers & Structures*, 2024, 297: 107342.
- [36] Kamalam G K, Rajasekar V. Digital twin meets artificial intelligence: AI-augmented industrial automation systems using intelligent digital twins[M]//*Digital Twin for Smart Manufacturing*. Amsterdam: Elsevier, 2023: 145–159.
- [37] Li X H, Chen B T, Fan J C, et al. Cloud-edge-end collaborative intelligent service computation offloading: A digital twin driven edge coalition approach for industrial IoT[J]. *IEEE Transactions on Network and Service Management*, 2024, 21(6): 6318–6330.
- [38] Maksimović M, Jokić S, Bošković M Č. Digital twins and 5G: Unlocking the potential in the energy sector[C]//*Proceedings of 23rd International Symposium INFOTEH-JAHORINA (INFOTEH)*. Piscataway, NJ: IEEE, 2024: 1–6.
- [39] Liu Z. A personality model of virtual characters[C]//*Proceedings of 7th World Congress on Intelligent Control and Automation*. Piscataway, NJ: IEEE, 2008: 2497–2500.
- [40] Cheng Y B, Wang Y Y. Evaluating the effect of outfit on personality perception in virtual characters[J]. *Virtual Worlds*, 2024, 3(1): 21–39.
- [41] Zhu X Y, Yu C, Huang D, et al. Beyond 3DMM: Learning to capture high-fidelity 3D face shape[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(2): 1442–1457.
- [42] Dou P F, Shah S K, Kakadiaris I A. End-to-end 3D face reconstruction with deep neural networks[C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway, NJ: IEEE, 2017: 1503–1512.
- [43] Shen K Y, Guo C, Kaufmann M, et al. X-avatar: Expressive human avatars[C]//*Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway, NJ: IEEE, 2023: 16911–16921.
- [44] Tan X, Chen J W, Liu H H, et al. NaturalSpeech: End-to-end text-to-speech synthesis with human-level quality[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(6): 4234–4245.
- [45] Sha F, Saul L K. Large margin hidden Markov models for automatic speech recognition[M]//*Advances in Neural Information Processing Systems 19*. Massachusetts: The MIT Press, 2007: 1249–1256.
- [46] Ruinskiy D, Lavner Y. An effective algorithm for automatic detection and exact demarcation of breath sounds in speech and song signals[J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2007, 15(3): 838–850.
- [47] Łajszczak M, Cámbara G, Li Y, et al. Base tts: Lessons from building a billion-parameter text-to-speech model on 100k hours of data[J]. *arXiv*, 2024: 2402. 08093.
- [48] Gibiansky A, Arik S, Diamos G, et al. Deep voice 2: Multi-speaker neural text-to-speech[J]. *Advances in neural information processing systems*, 2017: arXiv: 1705.08947.
- [49] Sabour S, Zhang W, Xiao X Y, et al. A chatbot for mental health support: Exploring the impact of Emohaa on reducing mental distress in China[J]. *Frontiers in Digital Health*, 2023, 5: 1133987.
- [50] Na H. CBT-LLM: A chinese large language model for cognitive behavioral therapy-based mental health question answering[C]//*In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation*. Torino: ELRA and ICCL, 2024: 2930–2940.
- [51] Liu C X, Xie Z Y, Zhao S R, et al. Speak from heart: An emotion-guided LLM-based multimodal method for emotional dialogue generation[C]//*Proceedings of the 2024 International Conference on Multimedia Retrieval*. New York: ACM, 2024: 533–542.
- [52] Llanes-Jurado J, Gómez-Zaragoza L, Minissi M E, et al. Developing conversational Virtual Humans for social emotion elicitation based on large language models[J]. *Expert Systems with Applications*, 2024, 246: 123261.
- [53] van der Struijk S, Huang H H, Mirzaei M S, et al. FACS-vatar: An open source modular framework for real-time FACS based facial animation[C]//*Proceedings of the 18th International Conference on Intelligent Virtual Agents*. New York: ACM, 2018: 159–164.
- [54] Garg K, Kello C T. Efficient Lévy walks in virtual human foraging[J]. *Scientific reports*, 2021, 11(1): 5242.
- [55] 腾讯新闻. 海蓝手语数字人出圈!山东一千河数字科技斩获2024中国正能量网络传播AI精品案例[EB/OL]. [2025-03-11]. <https://news.qq.com/rain/a/20250311A08KGB00>.
- [56] 中国政府网. 中国政府网“AI数字人”为你讲政策: 第一次写进2025《政府工作报告》[EB/OL]. [2025-03-07]. https://www.gov.cn/yaowen/shipin/202503/content_7011583.htm.
- [57] 搜狐科技. AI数字人视频生成方案: 提升广电媒体直播效率和智能化水平[EB/OL]. [2024-07-01]. https://it.sohu.com/a/789777193_120912206.
- [58] 微软公司. 小冰基于微软Azure向亚洲客户推出新一代AI数字员工解决方案[EB/OL]. [2023-07-24]. <https://news.microsoft.com/zh-cn/小冰基于微软azure向亚洲客户推出新一代ai数字员工解决方案>.
- [59] 澎湃新闻. AI赋能, 创新中华优秀传统文化的传播之道[EB/OL]. [2024-10-05]. https://www.thepaper.cn/newsdetail_forward_28942518.
- [60] 南都全国两会前方报道组. 带着专属“数字人”上两会能做什么?对话全国人大代表张敏[EB/OL]. [2025-03-24]. <https://news.qq.com/rain/a/20250311A08REQ00>.
- [61] 人民日报. 数字人主播“上岗”了[EB/OL]. [2025-02-28]. http://paper.people.com.cn/rmrhwb/pc/content/202502/28/content_30059410.html.
- [62] 科学网. 上海交大: “虚拟数字人”打造教育传播新范式[EB/OL]. [2023-12-08]. <https://news.sciencenet.cn/html-news/2023/12/513867.shtm>.
- [63] 搜狐网. 数字人直播助力教育创新, 风平智能与新东方联手开创新局![EB/OL]. [2023-08-02]. <https://www.sohu>.

- com/a/708386474_121060424.
- [64] 网易. AI数字人赋能教育行业智能课件录制, 快速生成网络课堂教学视频[EB/OL]. [2023-10-11]. <https://www.163.com/dy/article/IGPNQF85055300PO.html>.
- [65] 李杰. 元宇宙背景下《雪崩》的引文分析[J]. 科学观察, 2023, 18(1): 45-56.
- [66] Cline E. Ready player one[M]. New York: Ballantine Books, 2011.
- [67] Huang Y, Li Y J, Cai Z P. Security and privacy in metaverse: A comprehensive survey[J]. Big Data Mining and Analytics, 2023, 6(2): 234-247.
- [68] 中国信息通信研究院. 工业元宇宙白皮书[EB/OL]. [2023-03-24]. <https://aai-alliance.org/uploads/1/20230324/8c43ae95c9991a94d3b71e61f6d6ba8a.pdf>.
- [69] Ren F J. Construction of metaverse center based on advanced intelligence[EB/OL]. [2023-12-28]. <https://www.interpaper.org/2022-Annual-Report-EU-Academy-ofSciences.pdf>.
- [70] 邓佳文, 任福继. 2023年生成式AI大模型发展热点回眸[J]. 科技导报, 2024, 42(1): 266-285.
- [71] Miao Q, Li K, Quan J, et al. Advances in 4d generation: A survey[J]. arXiv, 2025: 2503. 14501.

Advances in digital twin technology under the “object–human–environment” interaction perspective

YOU Jiali¹, DENG Jiawen^{1*}, JIAO Ziyun¹, LUO Ali², SONG Yihan², QIU Bo³, REN Fuji^{1*}

1. School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

2. National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China

3. School of Intelligent Science and Technology, University of Science and Technology, Beijing 100083, China

Abstract With the convergence and innovation of emerging information technologies, Digital Twin (DT) technology has become a key enabler for digital transformation and the evolution of intelligent systems. It has been widely applied and received significant attention in fields such as industrial manufacturing, smart cities, and intelligent transportation. However, existing research on traditional DT technologies has predominantly focused on the modeling and analysis of physical entities ("objects"), with limited systematic integration of "human" and "environmental" factors. The lack of exploration into human–environment interactions makes it difficult for current digital twin frameworks to meet the advanced demands of complex intelligent systems for multi–level, comprehensive interaction capabilities. In view of this, this paper innovatively introduces the "object–human–environment" interactive vision and comprehensively and systematically analyzes the research frontiers and progress of digital twin technology from the three core dimensions of intelligent physical entity (object), intelligent individual (human), and virtual–real fusion environment (environment). Firstly, the paper analyzes the traditional digital twin technology system with "object" as the core and focuses on its theoretical origin, framework, and application. Secondly, it discusses the definition, development context, national policies, and core technologies of digital people driven by AI. Finally, expand the vision to the dimension of "environment" and explore the application practice of "environment" in multiple scenes of the meta–universe, deeply discuss the deep integration and interaction mechanism of the three elements of "object", "human" and "environment", reveal how the three interact and promote each other, and provide support for the construction of the meta–universe. Furthermore, this study discusses the current research challenges and future development trends of digital twin technology from the perspective of "Object–Human–Environment" interaction and proposes three key research directions: (1) developing an intelligent, multi–layered data fusion framework; (2) exploring AIGC–enabled intelligent virtual–real mapping and native virtual evolution; and (3) constructing novel virtual economy architectures and intelligent governance systems. The research outcomes provide both theoretical foundations and practical insights for building next–generation digital twin systems characterized by multi–agent collaborative perception, multimodal intelligent interaction, and closed–loop integration of virtual and real environments.

Keywords digit twin; object–human–environment interaction; digital human; virtual–physical fusion; metaverse ●



(责任编辑 王微)